THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Structural and functional studies of suppressor of copper sensitivity proteins from *Proteus mirabilis*

Emily Jane Furlong

Bachelor of Science (Honours)

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2018

Institute for Molecular Bioscience

Abstract

The correct formation of disulfide bonds is a crucial step in the folding of many virulence factors in bacteria. In many Gram-negative bacteria, redox machinery in the periplasm ensures this occurs efficiently and correctly for secreted and membrane proteins through the processes of disulfide bond oxidation and isomerisation. This machinery has been well characterised in *Escherichia coli* (strain K-12) where it consists of four key disulfide bond forming (Dsb) proteins, DsbA, DsbB, DsbC and DsbD. *E. coli* DsbA is the prototypical bacterial oxidase, which catalyses the formation of disulfide bonds in its protein substrates. When the active site cysteines of DsbA become reduced, DsbA needs to be reoxidised by its redox partner, DsbB, to continue to function. DsbA often forms disulfide bonds incorrectly and DsbC corrects this by isomerising or shuffling the disulfide bonds in the substrate protein. As DsbC is only active when its catalytic cysteines are reduced, it also has a redox partner, DsbD.

Over the past 27 years, Dsb-like proteins from various bacteria have been studied and it has become clear that there is great diversity in the number and type of proteins that contribute to the redox machinery in different bacteria. Some bacteria have multiple Dsb proteins and distinct Dsb-like systems that have specific roles. The suppressor of copper sensitivity (Scs) proteins form one such system. Originally identified in *Salmonella enterica* serovar Typhimurium, the *Scs* locus encodes four proteins (ScsA, ScsB, ScsC and ScsD). The *S.* Typhimurium locus was found to confer copper tolerance to a copper sensitive *E. coli* mutant, so it is thought that these proteins play a role in bacterial response to copper stress. Three Scs proteins (ScsB, ScsC and ScsD) have predicted thioredoxin folds and catalytic cysteines, similar to the Dsb proteins. This thesis focuses on the characterisation of the ScsC and ScsB proteins from the uropathogen *Proteus mirabilis*.

Prior to this work, ScsC had only been studied in two organisms, *S.* Typhimurium and *Caulobacter crescentus*. Chapter 2 is a *Nature Communications* paper that describes the genetic, biochemical and structural characterisation of *P. mirabilis* ScsC (PmScsC). This work revealed that PmScsC is a highly dynamic, trimeric disulfide isomerase that increases the swarming motility of *P. mirabilis* under copper stress. An N-terminally truncated PmScsC variant, PmScsCΔN, is monomeric and has little *in vitro* disulfide isomerase activity but is highly active as a dithiol oxidase. This variant could not complement the loss of ScsC in *P. mirabilis,* suggesting that PmScsC functions as a trimeric disulfide isomerase *in vivo*. The three crystal structures of PmScsC revealed that

the region responsible for the highly dynamic nature of the protein is an 11 amino acid flexible linker between the trimerisation domain and catalytic domain. Replacement of this flexible linker with a rigid linker revealed that it was necessary for the disulfide isomerase activity of PmScsC.

Chapter 3 describes the structural characterisation of the PmScsC variant, PmScsCΔN and the biochemical and structural characterisation of another variant PmScsCΔLinker. The structure of PmScsCΔN is very similar to the catalytic domains of the native PmScsC structures, suggesting that it is the lack of the trimerisation stem and thus change in oligomerisation state that converts it from an isomerase to an oxidase. To assess the role of the flexible linker in PmScsC, the linker residues (39-49) were deleted. This deletion mutant (PmScsCΔLinker) crystallised as a trimer but exists in different oligomeric states in solution and does not function as an oxidase or an isomerase. The characterisation of these variants adds to our understanding of the key structural features of PmScsC and their influence on the function of this trimeric disulfide isomerase. This chapter has been submitted to *Acta Crystallographica Section D* for publication.

Based on work performed in *C. crescentus*, we hypothesised that the N-terminal periplasmic domain of PmScsB (PmScsBα) is the redox partner of PmScsC. Chapter 4 is a *Journal of Biological Chemistry* publication that describes the experiments performed to confirm this hypothesis. This publication also reports the crystal structure of PmScsBα, (the first ScsBα structure to be solved) and a SANS derived model of the complex formed between PmScsC and PmScsBα. I found that PmScsBα consists of two immunoglobulin-like domains, both of which interact with one PmScsC protomer in the SANS model.

To gain more information about PmScsB, I wanted to produce the full-length transmembrane protein for further characterisation. Chapter 5 describes the work performed towards the design of robust expression and purification protocols for the production of PmScsB.

In summary, this thesis describes the characterisation of PmScsC, PmScsC mutants, PmScsBα and the interaction between PmScsC and PmScsBα. It also contains the steps taken toward the production of the full-length transmembrane protein PmScsB for future analysis. This work adds to our understanding about the Suppressor of copper sensitivity proteins and the diversity of redox machinery in bacteria.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications included in this thesis

Chapter 2

Furlong, E. J., Lo, A. W., Kurth, F., Premkumar, L., Totsika, M., Achard, M. E. S., Halili, M. A., Heras, B., Whitten, A. E., Choudhury, H. G., Schembri, M. A., and Martin, J. L. (2017) A shape-shifting redox foldase contributes to *Proteus mirabilis* copper resistance. *Nature Communications* **8**, 16065; doi: 10.1038/ncomms16065.

Chapter 4

Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E. and Martin, J. L. (2018) Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB. *Journal of Biological Chemistry* **293**, 5793-5805; doi: 10.1074/jbc.RA118.001860.

Submitted manuscripts included in this thesis

Chapter 3

Furlong, E. J., Kurth, F., Premkumar, L., Whitten, A. E., Martin, J. L. (2018) Engineered variants provide new insight into the structural properties important for activity of the highly dynamic, trimeric protein disulfide isomerase, PmScsC. Available on bioRxiv; doi: 10.1101/421420. Submitted to *Acta Crystallographica Section D*.

Other publications during candidature

Peer-reviewed papers

1) Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E. and Martin, J. L. (2018) Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB. *Journal of Biological Chemistry* **293**, 5793-5805; doi: 10.1074/jbc.RA118.001860.

2) Furlong, E. J., Lo, A. W., Kurth, F., Premkumar, L., Totsika, M., Achard, M. E. S., Halili, M. A., Heras, B., Whitten, A. E., Choudhury, H. G., Schembri, M. A., and Martin, J. L. (2017) A shape-shifting redox foldase contributes to *Proteus mirabilis* copper resistance. *Nature Communications* **8**, 16065; doi: 10.1038/ncomms16065.

3) Christensen, S., Groftehauge, M. K., Byriel, K., Huston, W. M., **Furlong, E.**, Heras, B., Martin, J. L., and McMahon, R. M. (2016) Structural and Biochemical Characterization of Chlamydia trachomatis DsbA Reveals a Cysteine-Rich and Weakly Oxidising Oxidoreductase. *PLoS One* **11**, e0168485.

Conference abstracts

1) Furlong, E. J., Kurth, F., Choudhury, H. G., Premkumar, L., Duff, A. P., Whitten, A. E. and Martin, J. L. Studies of the trimeric disulfide isomerase PmScsC and its redox partner PmScsBα. Invited speaker in the “Hot Structures” microsymposium, AsCA 2018/Crystal 32, 2-5 December 2018, Auckland, New Zealand.

2) Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E. and Martin, J. L. Structural and functional analysis of two *Proteus mirabilis* copper resistance proteins

reveals an unusual redox relay system. Selected presentation in the “Rising Star” plenary session, Crystal 31, 3-7 December 2017, Bunker Bay, Western Australia.

3) Furlong, E. J., Kurth, F. Premkumar, L., Choudhury, H. G. and Martin, J. L. *P. mirabilis* ScsC: A trimeric, shape-shifting disulfide isomerase. Poster presentation, International School of Crystallography, 50th Course: Integrative Structural Biology, 2-11 June 2017, Erice, Sicily.

4) Furlong, E. J., Kurth, F., Premkumar, L., Choudhury, H. G. and Martin, J. L. *P. mirabilis* ScsC: A trimeric, shape-shifting disulfide isomerase. Selected oral presentation in the “Trafficking, Pathogens and Parasites” colloquium session, ComBio 2016, 3-7 October 2016, Brisbane, Australia.

5) Furlong, E. J., Kurth, F. Premkumar, L., Choudhury, H. G. and Martin, J. L. *Proteus mirabilis* disulfide isomerase - structure and function. Poster presentation, 41st Lorne Conference on Protein Structure and Function, 7-12 February 2016, Lorne, Victoria, Australia.

Contributions by others to the thesis

Chapter 2

- Dr Alvin Lo constructed the *P. mirabilis* mutants, generated the complementation plasmids, performed the motility assays and examined ScsC expression by western blotting and contributed to the interpretation of these results. He also contributed to the writing of the methods section in the manuscript.
- Dr Fabian Kurth created the PmScsC and PmScsCΔN constructs for protein purification and expression, optimised PmScsC expression and purification protocols and produced protein used in the experiments. He also performed and analysed data for the biochemical characterisation of PmScsC and PmScsCΔN, evaluated the *P. mirabilis* clinical strains for ScsC, crystallised one of the three crystal structures (compact PmScsC) and was involved in collecting and processing data to solve the compact PmScsC crystal structure. He also contributed to the writing of the manuscript.
- Dr Lakshmanane Premkumar collected and processed data to solve the compact PmScsC crystal structure and completed its refinement.

- Dr Makrina Totsika gave advice and assistance in evaluating the *P. mirabilis* clinical strains for ScsC.
- Dr Maud Achard tested the growth of wild-type and mutant *P. mirabilis* strains in the presence and absence of copper and analysed these results.
- Dr Maria Halili was involved in performing the *in vitro* biochemical characterisation of native PmScsC.
- Dr Begoña Heras gave advice and assistance for the *in vitro* biochemical characterisation of native PmScsC.
- Dr Hassanul Choudhury assisted with collecting and processing data to solve the extended and transitional PmScsC crystal structures and helped with the initial refinement of the transitional structure. He also assisted in the analysis and interpretation of the crystal structures.
- Dr Andrew Whitten designed the SAXS experiments, collected, analysed and modeled the SAXS data and interpreted these results. He also contributed to the writing of the methods section in the manuscript.
- Professor Mark Schembri designed and directed the project, contributed to the interpretation of results and editing of the manuscript.
- Professor Jennifer Martin designed and directed the project, contributed to the refinement of the crystal structures and writing of the manuscript.

Chapter 3

- Dr Fabian Kurth crystallised PmScsCΔN, processed the diffraction data and solved and refined the crystal structure.
- Dr Lakshmanane Premkumar assisted with the data collection, data processing, structure solution and refinement of PmScsCΔN.
- Dr Andrew Whitten assisted with the design of the SAXS experiments and data collection and analysed and modeled the SAXS data and interpreted these results. He also contributed to the writing of the SAXS methods and results sections of the manuscript.
- Professor Jennifer Martin supervised the project and contributed to its design, assisted with the refinement of the PmScsCΔLinker crystal structure and critically revised the manuscript.

Chapter 4

- Dr Hassanul Choudhury contributed to solving the PmScsBα crystal structure and was involved in its interpretation.
- Dr Fabian Kurth designed and created the PmScsBα construct, optimised the PmScsBα expression and purification protocols and gained initial crystallisation hits.
- Dr Anthony Duff expressed the deuterated PmScsBα needed for SANS analysis and wrote the deuterated expression methods section of the manuscript.
- Dr Andrew Whitten designed and performed the SAXS and SANS experiments. Andrew also processed and analysed the SAS data, produced the PmScsC-PmScsBα model and wrote the SAS methods section of the manuscript.
- Professor Jennifer Martin supervised the project and contributed to its design and substantially edited the manuscript.

Chapter 5

- Dr Gordon King performed the mass spectrometry experiments and analysed and interpreted the results.
- Professor Jennifer Martin supervised the project.

Statement of parts of the thesis submitted to qualify for the award of another degree

No works submitted towards another degree have been included in this thesis.

Research Involving Human or Animal Subjects

No animal or human subjects were involved in this research.

Acknowledgements

Firstly, I would like to thank my supervisor, Jenny Martin. I first met Jenny in 2011 during the first year of my Bachelor of Science at UQ and over the past seven years she has been more than just a supervisor to me. She has given me opportunities that I could never have even dreamt of, encouraged me to step out of my comfort zone, supported me through difficult times and sponsored my endeavour to a pursue a career in research. I am extremely grateful to have worked in Jenny's lab over the past three and a half years. I would also like to thank Brett Collins for his supervision and support during my candidature and Hassan Choudhury for teaching me crystallography along with some important life lessons.

Of course, a massive thank you goes to the Institute for Molecular Bioscience postgraduate coordinator, Amanda Carozzi, who is such supportive, kind and generous person. I feel very lucky to have completed my PhD in an institute where there is so much support for students and research. Half way through my PhD the Martin lab moved to the Griffith Institute for Drug Discovery, so I would like to thank everyone there who welcomed us and made the transition a little smoother. I would also like to acknowledge Nick Hamilton, Luke Guddat and Michael Landsberg for being on my thesis committee. Throughout my PhD, I also had access to UQROCX, ANSTO and the Australian Synchrotron; without these facilities and the people that run them I would have very little data to present in my thesis.

If there is one thing I have learnt during my PhD it's that the people you work with can have a huge influence on your enjoyment of work, and life in general. So thank you to the Martin lab members, past and present, for being such wonderful people. Particularly, I would like to thank Russell Jarrott for letting me "borrow" his reagents, helping me with pretty much any lab-related problem and keeping the lab in working order. Thank you to Róisín McMahon, Camila Cotrim and Gordon King for helpful discussions about my experiments. Thanks to Martin lab long-term collaborator and former member, Andrew Whitten, for assisting me with small-angle scattering experiments and former Martin lab PhD student, Fabian Kurth, for leaving records so good that I could easily continue on from his previous work.

I spent six months of my PhD at the Sir William Dunn School of Pathology, The University of Oxford learning cryo-electron microscopy with Susan Lea. Though none of that work

has made it into my thesis, I would like to thank the whole Lea lab and Susan for hosting me, as well as all the people I met, who made those months some of the best in my life so far.

A big thank you to all of my friends and fellow PhD students that have, at some point in the past three and a half years, listened to me whinge or asked me how things were going. I'd particularly like to thank Emma Livingstone, Miranda Pitt, Sarah Piper, James Hill, Claudia Stocks, Chaseleigh Bradley, Brandon Binnie, Jessica Chitty and Jake Parker for being there whenever I needed to chat. Thanks also to my past and present housemates, Carson, Shweta, Shea and Kim, for keeping me grounded and taking a genuine interest in my work and life.

Finally, I am so grateful to my family for their never-ending support and encouragement. Thank you to my parents and grandparents, particularly, for always believing in me and devoting so much of their time, energy and love to me from day dot. I am incredibly lucky!

Financial support

This research was supported by an Australian Government Research Training Program Scholarship and an Institute for Molecular Bioscience Research Advancement Award.

Throughout my candidature I have received awards to attend conferences and further my career development. These are listed below.

- Australian Microscopy and Microanalysis Society travel grant to attend CryOz18 in Melbourne (300 AUD)
- Society of Crystallographers in Australian and New Zealand Maslen scholarship to attend the Crystal 31 conference in Perth (850 AUD)
- 2017 International School of Crystallography (Erice, Sicily) registration grant (1200 EUR)
- UQ Graduate School International Travel Award (5000 AUD) to spend six months being trained in cryoEM at The University of Oxford, UK, with Professor Susan Lea
- Society of Crystallographers in Australian and New Zealand Maslen scholarship (600 AUD) to attend the Australasian Crystallography School, held in Melbourne 2016

Keywords

Redox proteins, copper, bacteria, disulfide isomerase, oxidative folding, disulfide bond, *Proteus mirabilis*

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 060112, Structural Biology, 60%

ANZSRC code: 030403, Characterisation of Biological Macromolecules, 40%

Fields of Research (FoR) Classification

FoR code: 0601, Biochemistry and Cell Biology, 60%

FoR code: 0304, Medicinal and Biomolecular Chemistry, 40%

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMS	4-acetamide-4'-maleimidylstilbene-2-2'-disulfonate
ANSTO	Australian Nuclear Science and Technology Organisation
APBS	Adaptive poisson-boltzmann solver
BSA	Bovine serum albumin
Cc	*Caulobacter crescentus*
CCD	Charge-coupled device
cCMP	Cytidine 3',-5'-cyclic monophosphate
CCTOP	Constrained Consensus Topology server
CD	Circular Dichroism
CMC	Critical micelle concentration
CPM	N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]-maleimide
CryoEM	Cryogenic electron microscopy
CV	Column volume
DDM	*n*-Dodecyl-β-D-maltopyranoside
DM	*n*-Decyl ß-maltopyranoside
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOTA	1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid
Dsb	Disulfide bond forming protein
DTNB	5,5-dithio-bis-(2-nitrobenzoic acid)
DTSSP	Dithiobis(sulfosuccinimidylpropionate)
DTT	1,4-Dithiothreitol
Ec	*Escherichia coli*
EDTA	2,2',2'',2'''-(Ethane-1,2-diyldinitrilo)tetraacetic acid
EMBO	European Molecular Biology Organisation
EM	Electron microscopy
FOM	Figure of merit
FPLC	Fast protein liquid chromatography
FSEC	Fluorescence-detection size exclusion chromatography
GAE	Gamma-adaptin ear
GFP	Green-fluorescent protein
GGA	Golgi-localised gamma-ear-containing ADP ribosylation factor-binding
GSSG/GSH	Glutathione oxidised/reduced
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid

ID	Identifier
IMAC	Immobilised-metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
LCP	Lipidic cubic phase
LDAO	*N,N*-Dimethyldodecylamine *N*-oxide
LMNG	Lauryl Maltose Neopentyl Glycol
MALDI	Matrix assisted laser desorption/ionisation
MALLS	Multiangle laser light scattering
MCA	methoxy-coumarin amide
MES	2-(N-morpholino)ethanesulfonic acid
MOPS	3-morpholinopropane-1-sulfonic acid
MR	Molecular replacement
NA	Not applicable
NADPH	Nicotinamide adenine dinucleotide phosphate
NM	*n*-Nonyl-β-D-maltopyranoside
NMR	Nuclear magnetic resonance
OD	Optical density
OG	*n*-Octyl-β-D-glucopyranoside
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PDI	Protein disulfide isomerase
PEG	Polyethylene glycol
Pm	*Proteus mirabilis*
RFU	Relative fluorescence unit
RHP	Rigid helical peptide
RMS	Root mean squared
RMSD	Root mean squared deviation
RNA	Ribonucleic acid
RNase A	Ribonuclease A
ROS	Reactive oxygen species
SAD	Single anomalous diffraction
SANS	Small angle neutron scattering

SAS	Small angle scattering
SASBDB	Small angle scattering biological data bank
SAXS	Small angle X-ray scattering
ScRNase A	Scrambled ribonuclease A
Scs	Suppressor of copper sensitivity
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SeMet	Selenomethionine
Sf	Superfolder
SOC	Super optimal broth with catabolite repression
St	*Salmonella enterica* serovar Typhimurium
TB	Terrific broth
TCA	Trichloroacetic acid
TEV	Tobacco etch virus
TLS	Translation-Libration-Screw refinement
TM	Transmembrane helix
TOF	Time of flight
TR	Thioredoxin reductase
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
Triton-X	Polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether
TRX	Thioredoxin
UDM	*n*-Undecyl-β-D-maltopyranoside
UPEC	Uropathogenic *Escherichia coli*
UQ ROCX	University of Queensland Remote Operation Crystallisation and X-ray
UTI	Urinary tract infection
UV	Ultraviolet
WAXS	Wide angle X-ray scattering
WEHI	Walter and Eliza Hall Institute
WT	Wild-type

Chapter 1

1. Introduction

This thesis focuses on bacterial redox proteins, specifically the suppressor of copper sensitivity proteins from *Proteus mirabilis*. This chapter begins by introducing the related disulfide bond forming proteins from *Escherichia coli* and summarising the conservation of these proteins across different bacterial species. The antibacterial action of copper and the response of bacteria to copper stress are then discussed before the *P. mirabilis* suppressor of copper sensitivity proteins are introduced. This chapter also provides the significance and aims of the research in this thesis and outlines how the subsequent chapters address each aim.

1.1 Disulfide bond forming proteins

The formation of disulfide bonds is an important step in the folding of many proteins. These covalent bonds, which are formed between the sulfur atoms of two cysteines, add stability to many folded proteins, allowing them to maintain their active, three-dimensional structure. In bacteria, proteins that require structural disulfide bonds are often virulence factors that are found in the periplasm or are transported into the harsh extracellular environment (1,2). Many Gram-negative bacteria have periplasmic machinery that is dedicated to forming disulfide bonds in secreted proteins. This machinery comprises a family of proteins, aptly named the Disulfide bond forming (Dsb) proteins. Dsb proteins have at least two redox active cysteines that cycle between the disulfide-bonded and reduced thiol states to perform their catalytic function. The prototypical *E. coli* Dsb (EcDsb) machinery consists of four proteins; EcDsbA, EcDsbB, EcDsbC and EcDsbD. EcDsbA and EcDsbB are involved in the oxidation step of disulfide bond formation (1,3), whereas EcDsbC and EcDsbD are involved in the isomerisation of incorrect disulfide bonds (4-6) (Figure 1.1).

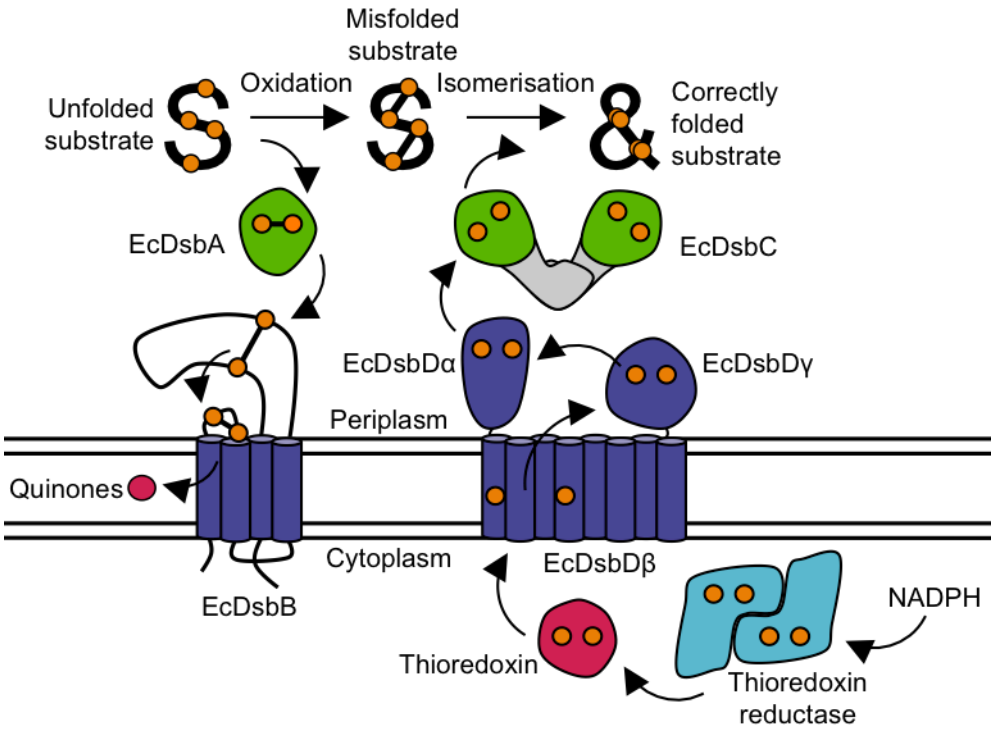


Figure 1.1: Overview of the prototypical* E. coli *Dsb system. *EcDsbA and EcDsbB, along with quinones, make up the disulfide bond oxidation pathway, responsible for introducing disulfide bonds into unfolded proteins. EcDsbC and EcDsbD are the key proteins in the isomerisation pathway, which rearranges incorrect disulfide bonds in substrate proteins. The isomerisation pathway relies on a reduction cascade that originates in the cytoplasm with NADPH. The curved arrows in the diagram indicate the flow of electrons throughout the pathways.*

1.1.1 *E. coli* DsbA

E. coli DsbA (EcDsbA) functions as the oxidase in the Dsb pathway, catalysing the formation of disulfide bonds in substrate proteins. The crystal structure of EcDsbA was published in 1993 (7) and shows that this monomeric, globular protein contains a thioredoxin fold. This fold (green in Figure 1.2A,B) consists of two motifs, βαβ and ββα, with a third α-helix situated between them (Figure 1.2A and B). EcDsbA also harbours an α-helical bundle between the ββα motif and the connecting α-helix of the thioredoxin fold. The redox active CXXC motif is positioned at the end of α-helix 1, and next to this site is a hydrophobic groove shown to be involved in the binding of other proteins (Figure 1.2C). The more N-terminal of the two catalytic cysteines (Cys30) is accessible to substrate proteins, whereas the second cysteine (Cys33) is buried in the protein. The pK_a of Cys30 is very low, approximately 3.5 (8), so the reactive thiolate anion form of its side chain is favoured at physiological pH 7. The formation of an intramolecular disulfide bond between

the catalytic cysteines of EcDsbA destabilises the protein (9), thus making the transfer of this disulfide bond favourable. This disulfide instability gives EcDsbA its highly oxidising nature (redox potential -122 mV (9,10)).

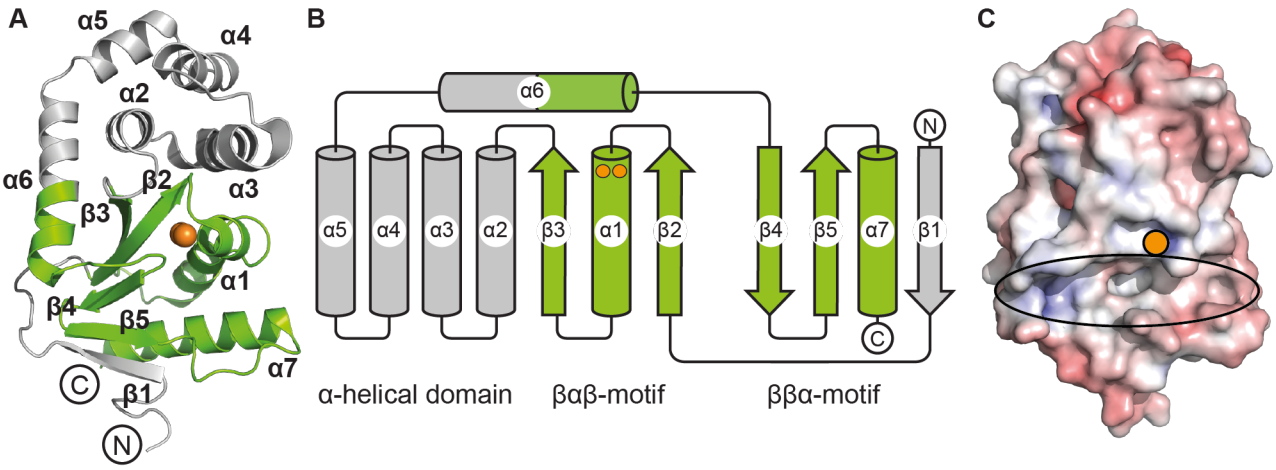


Figure 1.2: Structural features of EcDsbA. A. *Structure of EcDsbA (PDB ID: 1DSB) shown in cartoon representation.* ***B.*** *Topology diagram of EcDsbA. In both* ***A*** *and* ***B*** *the thioredoxin fold is coloured green, with the position of the catalytic cysteines shown as orange circles/spheres. The inserted helical domain and N-terminal β-strand are coloured grey.* ***C.*** *Electrostatic surface potential of EcDsbA. Electrostatic calculations were performed using APBS (11) and contoured to -7.5 and +7.5 k_BT/e. Positively charged surfaces are shown in blue, negatively charged surfaces are red and hydrophobic surfaces are white. An orange circle represents the position of the catalytic cysteines; the hydrophobic groove is surrounded by an oval.*

EcDsbA-catalysed disulfide bond formation is proposed to occur via two bimolecular nucleophilic substitution (S_N2) reactions (Figure 1.3) (12,13). In the first reaction, a thiolate anion of a cysteine residue in the unfolded substrate attacks Cys30 of oxidised EcDsbA, forming an intermolecular disulfide bond and releasing Cys33 as a thiol group (14,15). The intermolecular disulfide bond is then resolved in the second reaction, by the attack from another free thiolate anion in the substrate protein, resulting in disulfide exchange between EcDsbA and the substrate (15). A number of periplasmic proteins depend on EcDsbA for the formation of structural disulfide bonds (16). These include virulence factors such as FlgI, which forms part of the flagella motor (17) and EscC, which is an outer membrane secretin of the type III secretion system in enteropathogenic *E. coli* (18). Deletion of *dsbA* in *E. coli* results in the reduced production of properly folded virulence factors (1,17).

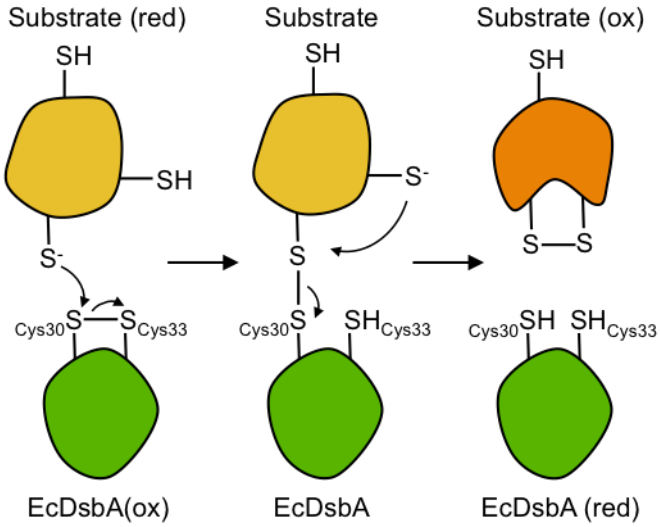


Figure 1.3: Catalytic action of EcDsbA. *The mechanism of EcDsbA-catalysed disulfide bond formation proceeds via two S_N2 reactions and results in the formation of a disulfide bond in the substrate protein and reduction of the catalytic cysteines in EcDsbA. EcDsbA is shown in green and the unfolded substrate is coloured yellow. The disulfide bonded or folded substrate is coloured orange.*

1.1.2 *E. coli* DsbB

After oxidising two thiol groups of a substrate protein, the catalytic cysteines of EcDsbA become reduced. To continue the cycle of disulfide bond formation in the periplasm, the catalytic cysteines of EcDsbA need to be oxidised (converted to the disulfide bonded state). As EcDsbA is more stable in the reduced form (9), a redox partner is required to oxidise EcDsbA and reactivate the enzyme. This redox partner is EcDsbB (3). EcDsbB is an inner membrane protein consisting of four transmembrane helices (Figure 1.4A) (19). The N- and C-termini are located in the cytoplasm and the protein has two periplasmic loops. The second periplasmic loop is much larger than the first and also contains a short α-helix (α4) that is associated with the membrane. Both loops contain an essential redox active pair of cysteines that need to be oxidised for EcDsbB activity (19). The second pair, Cys104 and Cys130, are needed for the interaction with EcDsbA (20,21), whereas the first pair, Cys41 and Cys44 are involved in the interaction with quinones, such as ubiquinone and menaquinone (22,23). These quinones are essential for EcDsbB catalytic function as they provide the driving force for EcDsbA oxidation by passing electrons obtained from EcDsbA onto electron acceptors involved in respiration, maintaining EcDsbB in the oxidised state (24).

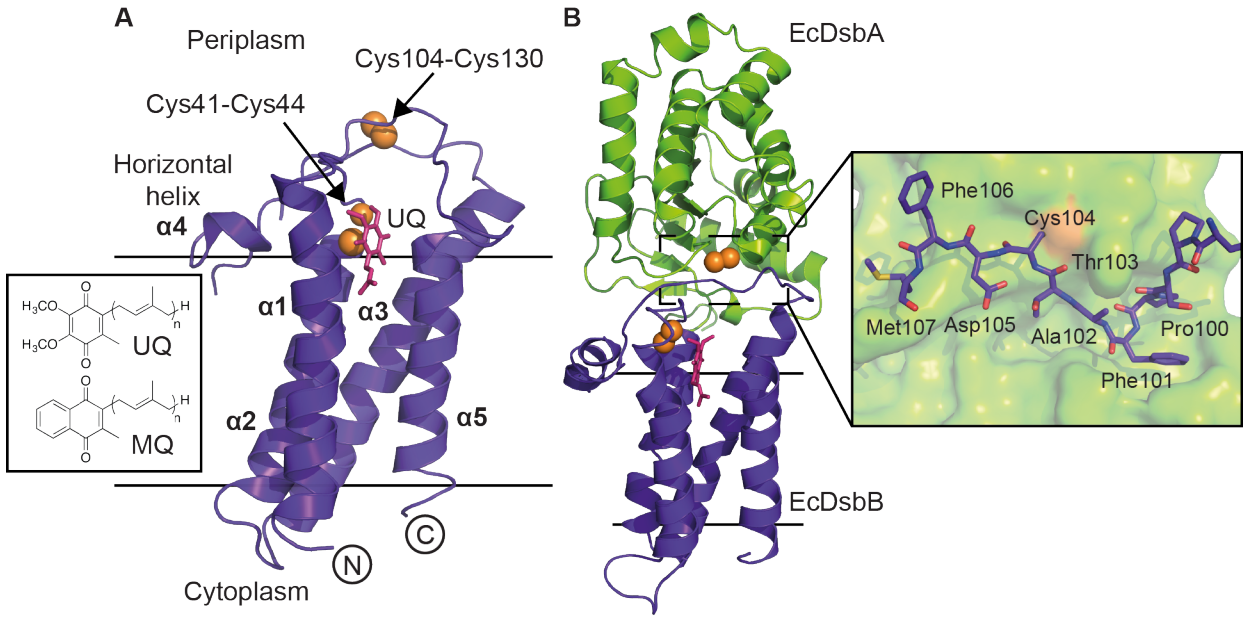


Figure 1.4: Structure of EcDsbB and the EcDsbA-EcDsbB interaction. A. *Crystal structure of the complex between antigen-binding fragment (not shown) and EcDsbB Cys41Ser mutant (PDB ID: 2ZUQ). EcDsbB is coloured purple, the positions of the catalytic cysteines are shown by orange spheres and the part of the ubiquinone molecule resolved in the structure is shown as magenta sticks. The chemical structures of ubiquinone (UQ) and menaquinone (MQ) are shown.* ***B.*** *Crystal structure of the complex between EcDsbA and EcDsbB (PDB ID: 2HI7). The colours in* ***A*** *are maintained except in* ***B*** *the orange spheres represent the disulfide bonded cysteines. The inset shows how the second periplasmic loop of EcDsbB binds into the hydrophobic groove of EcDsbA.*

The interaction between EcDsbA and EcDsbB has been well characterised. A crystal structure of the interaction was obtained by mutating Cys33 of EcDsbA to Ala and Cys130 of EcDsbB to Ser, which allowed the formation of an intermolecular disulfide bond between Cys30 (EcDsbA) and Cys104 (EcDsbB), resulting in a stable complex (Figure 1.4B) (25). The structure shows that the second periplasmic loop of EcDsbB binds into the hydrophobic groove close to the EcDsbA catalytic cysteines. There are two proposed mechanisms for the regeneration of EcDsbA by EcDsbB and both involve a series of disulfide exchange reactions and large conformational changes in EcDsbB that reposition the catalytic cysteine residues throughout the catalytic cycle (Figure 1.5) (26,27). For both mechanisms, the first step involves the attack of the reactive thiolate anion of Cys30 in EcDsbA on the Cys104-Cys130 disulfide bond in EcDsbB. This results in the formation of an intermolecular disulfide bond between Cys30 (EcDsbA) and Cys104 (EcDsbB) and the release of Cys130 (EcDsbB). The second and third steps differ between the reaction

mechanisms. In the “concerted” or “slow” model, the Cys130 thiolate anion attacks the Cys41-Cys44 disulfide bond of EcDsbB, forming an inter-loop disulfide between Cys130 and Cys41 (2a, Figure 1.5). The free Cys44 then forms a charge transfer intermediate with quinones, which gives EcDsbB a purple colour (28,29). Cys33 of EcDsbA then attacks the intermolecular EcDsbA-EcDsbB disulfide, resulting in the reoxidation of the EcDsbA cysteines and release of Cys104 (EcDsbB) (3a, Figure 1.5). In the alternate or “rapid” model essentially these two steps happen the other way round, that is, oxidised EcDsbA is released before Cys130 of EcDsbB attacks the Cys41-Cys44 disulfide bond (2b and 3b, Figure 1.5). In this model, it is thought that when EcDsbA binds EcDsbB a conformational change occurs in EcDsbB, distancing Cys130 so it is unable to catalyse the backward resolution of the EcDsbA-EcDsbB disulfide bond (25). It is also proposed that free Cys130 (EcDsbB) interacts with a loop on EcDsbA, prohibiting Cys130 (EcDsbB) from interacting with the Cys41-Cys44 disulfide, when EcDsbA is bound (26). Nevertheless, the fourth step in both pathways is the resolution of the Cys130-Cys41 disulfide by the attack of the Cys104 thiolate anion, resulting in the reformation of the Cys104-Cys130 bond (Figure 1.5). The Cys41 thiolate anion then attacks the Cys44-quinone complex, reforming the Cys41-Cys44 disulfide bond and releasing the now reduced quinone.

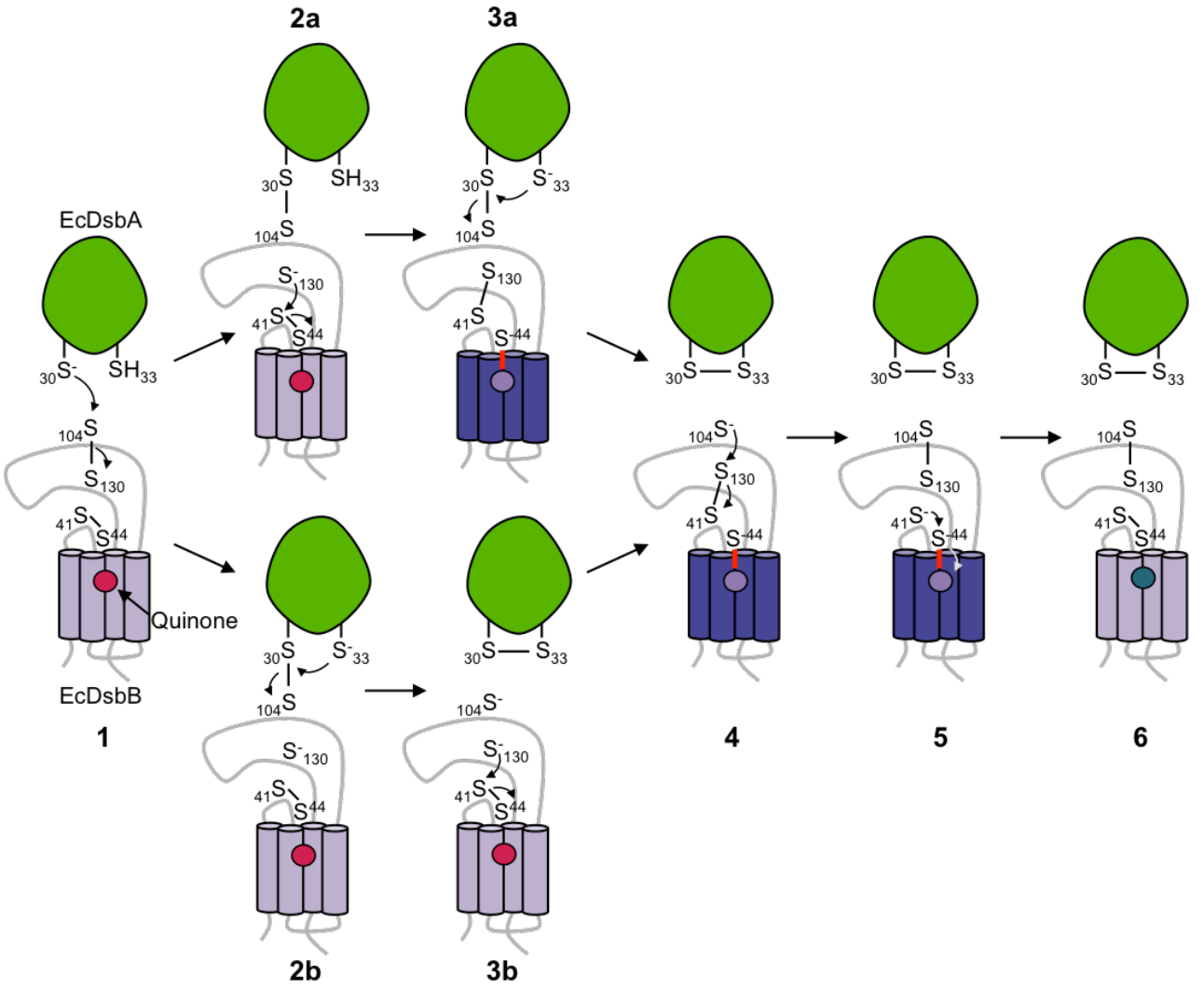


Figure 1.5: Proposed mechanisms for the reactivation of EcDsbA by EcDsbB. *The two proposed mechanisms for the EcDsbB-catalysed oxidation of EcDsbA differ at steps 2 and 3. Throughout the schematic, EcDsbA is shown in green and EcDsbB is shown in both light and dark purple. The colour changes of the quinone (circle) represent the different states of the electron acceptor throughout the pathway. The quinone is coloured magenta when associated with EcDsbB (light purple), light purple when it forms a charge transfer complex (represented by the red line) with EcDsbB (which changes to dark purple to represent the colour change) and dark turquoise once the electrons have been accepted. The curved arrows represent the flow of electrons.*

1.1.3 Inhibition of the oxidative folding pathway as an antibacterial strategy

Both EcDsbA and EcDsbB are considered targets for antibacterial drug discovery because of the important role they play in the formation of bacterial virulence factors. Further, a *dsbAB* knock-out strain of uropathogenic *E. coli* has reduced fitness in a mouse model of infection (30). The targeting of bacterial virulence as opposed to bacterial viability is an

emerging approach that is proposed to overcome the selective pressure for resistance that plagues the use of traditional antibiotics. Drug discovery targeting protein-protein interactions can be challenging, as there are generally no pockets on the interaction surface where a small molecule can easily bind. Nevertheless, fragment based drug discovery has been used to identify inhibitors that bind to EcDsbA (31) and peptide-based inhibitors have also been designed to bind to the reactive cysteine and into the hydrophobic groove of EcDsbA (32,33). EcDsbB has also been validated as an antivirulence target and high-throughput screening and fragment based drug discovery have been used to identify EcDsbB inhibitors (34-36). This has resulted in inhibitors that compete with quinone and or EcDsbA binding (34,35). Further work is needed to increase the potency of EcDsbA and EcDsbB inhibitors and research is now being undertaken to develop inhibitors of DsbA and DsbB from other bacteria (37,38).

1.1.4 *E. coli* DsbC

EcDsbA typically introduces disulfide bonds between cysteine residues into substrate proteins in a sequential manner along the unfolded polypeptide chain, as it is translocated into the periplasm (15). This allows for the correct folding of proteins that have disulfide bonds between consecutive cysteines, but results in the misfolding of substrates that need a different configuration of disulfide bonds, such as RNase I (17). Disulfide isomerases are able to shuffle disulfide bonds in a protein until it is correctly folded and are required for the folding of many proteins. EcDsbC is the prototypical bacterial disulfide isomerase and is a globular, V-shaped, homodimeric, periplasmic protein (Figure 1.6A) (4,39,40).

Each EcDsbC protomer contains a catalytic DsbA-like domain, comprised of a thioredoxin fold harbouring the redox active CXXC motif and an inserted helical-bundle consisting of only two helices (α4-5 in EcDsbA are replaced by a loop in EcDsbC) (Figure 1.6C) (39). Like EcDsbA, the N-terminal cysteine of EcDsbC (Cys98) is solvent accessible and acidic, with a pKa of 4.3 (41). It is expected that the catalytic cysteines of EcDsbC are also stabilised in the reduced form by a hydrogen-bonding network (39) and the redox potential of EcDsbC, at -130 mV (42), is only slightly more reducing than that of EcDsbA (-122 mV) (9,10). Dimerisation of EcDsbC is mediated by an N-terminal domain in each protomer that consists of a single N-terminal α-helix followed by 5 β-strands (39). The first four of these β-strands form a β-sheet with the fifth β-strand of the other EcDsbC protomer (Figure 1.6B). This intermolecular β-sheet holds the EcDsbC protomers together and dimerisation in this way is essential for the isomerase activity of EcDsbC (41,43). As the catalytic

domains of EcDsbA and EcDsbC are very similar, it is the dimerisation of EcDsbC that prevents cross talk between the Dsb oxidation and isomerisation pathways. Dimeric EcDsbC is required to interact with its redox partner EcDsbDα (44) and EcDsbB can only oxidise monomeric mutants of EcDsbC (45). The hydrophobic dimerisation cleft of EcDsbC is also predicted to be important for the interaction with substrates (Figure 1.6D) (39). Misfolded proteins typically have exposed hydrophobic patches, which are thought to be detected by EcDsbC and shielded by the enzyme's hydrophobic cleft whilst the disulfide bonds are shuffled by the catalytic cysteines (39). EcDsbC also functions as a chaperone (46) and the hydrophobic dimerisation cleft is essential for this activity (39,41).

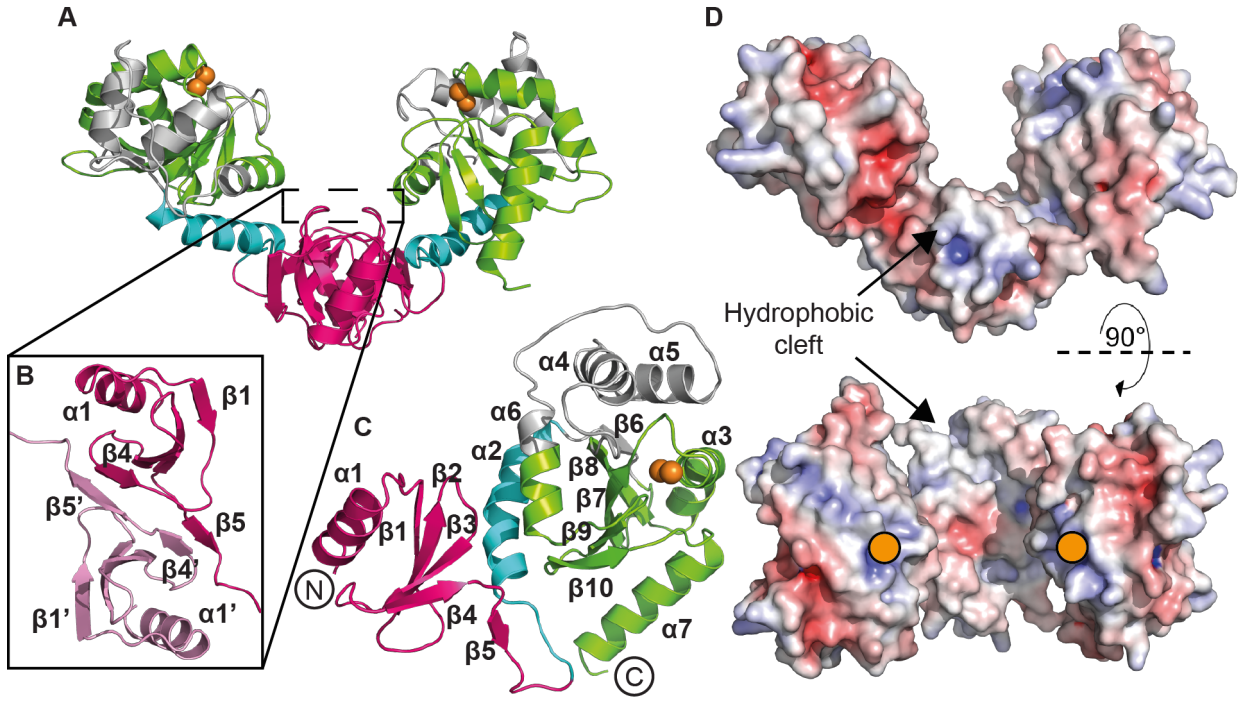


Figure 1.6: EcDsbC structure. A. *Dimeric EcDsbC with the dimerisation domain coloured magenta, the hinged linker helix shown in cyan and the thioredoxin fold coloured green with the catalytic cysteines represented by orange spheres. Other features, including the inserted helical domain, are grey.* ***B.*** *The dimerisation interface of EcDsbC consists of an intermolecular β-sheet, formed by β1-4 in one protomer and β5 in another.* ***C.*** *Structure of an EcDsbC protomer, coloured as in* ***A*** *with the secondary structure elements and the N- and C-termini labelled.* ***D.*** *Electrostatic surface potential of dimeric EcDsbC. Electrostatic calculations were performed using APBS (11) and contoured to -7.5 and +7.5 k_BT/e. Positively charged surfaces are shown in blue, negatively charged surfaces are red and hydrophobic surfaces are white. An orange circle represents the position of the catalytic cysteines.*

The catalytic and dimerisation domains of each EcDsbC protomer are connected by a hinged linker helix (39). This hinged linker gives the isomerase conformational flexibility, allowing the catalytic domains to move relative to each other. This was evidenced by the conformational change observed between the EcDsbC crystal structure (PDB ID: 1EEJ) (39) and that of the complex between EcDsbC and its redox partner EcDsbDα (PDB ID: 1JZD) (47). In the structure of the complex the distance between the two catalytic sites is ~29 Å, compared to ~38 Å in the EcDsbC structure (47). This flexibility is thought to facilitate the disulfide isomerase and chaperone activity of EcDsbC (39,47), by allowing the bound misfolded substrates to sample many different shapes. EcDsbC would need to accommodate substrate conformational changes and enable repositioning of the substrate's cysteines, allowing the correct disulfide bonds to be formed. The EcDsbC flexibility would also allow the enzyme to interact with a variety of different substrates.

In contrast to EcDsbA, EcDsbC functions in the reduced form, with the free thiolate anion of the reactive cysteine (Cys98) attacking disulfide bonds in substrate proteins (Figure 1.7) (48). Once the EcDsbC-substrate intermediate has formed there are two ways in which the reaction can proceed and both have been confirmed by experimental evidence (48-50). The first pathway involves the intermolecular disulfide being attacked by another cysteine in the substrate protein, resulting in the isomerisation of the substrate's disulfide bonds. In the second pathway, the intermolecular disulfide bond is resolved by Cys101 of EcDsbC, resulting in oxidised EcDsbC and the reduced substrate disulfide (51).

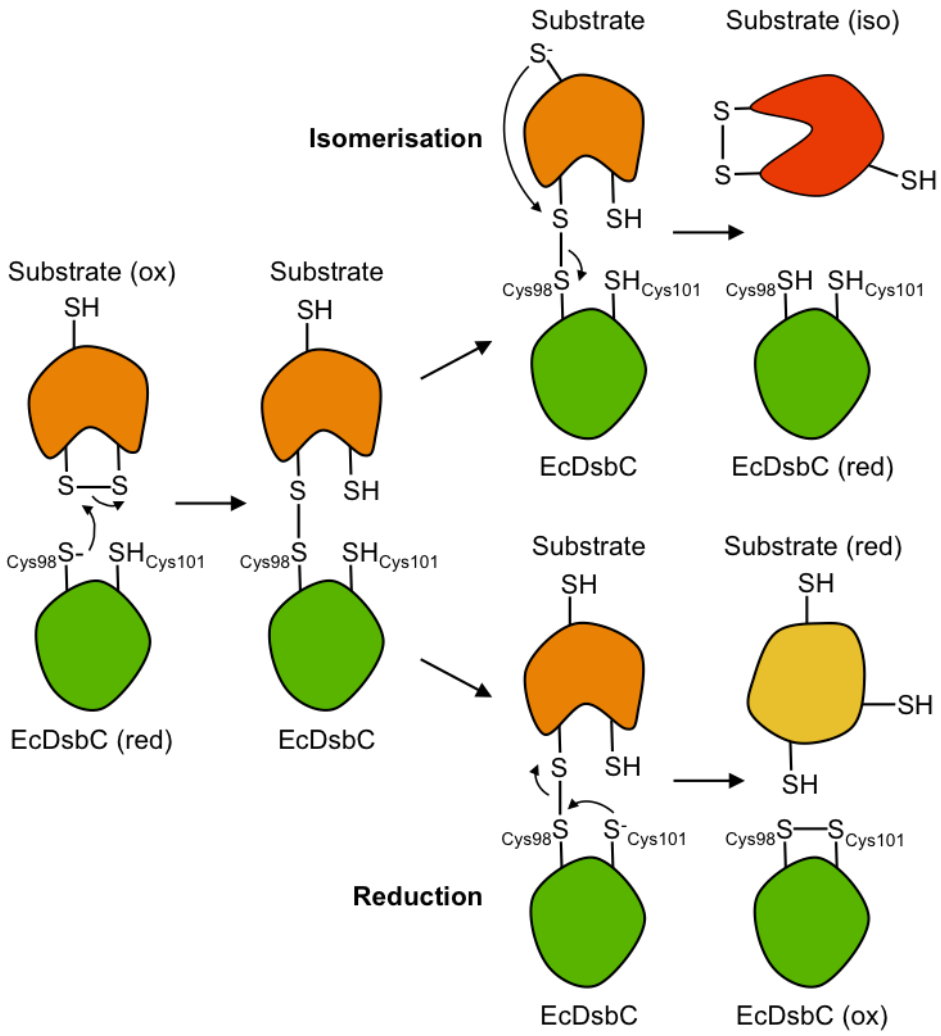


Figure 1.7: EcDsbC mechanism of action. *EcDsbC (green) functions in the reduced form and can either isomerise or reduce the disulfide bonds of substrate proteins. In the diagram, the substrate changes colour and shape depending on whether it is oxidised (orange or red), reduced (yellow) or the disulfide bonds are isomerised (orange vs red).*

Only a handful of EcDsbC substrates have been identified, including the penicillin-insensitive murein endopeptidase MepA, an endonuclease End1 and the ribonuclease RNase I (17,52). Under normal growth conditions, deletion of *dsbC* in *E. coli* has no clear phenotype (53) however the deletion strain does show reduced viability when exposed to excess copper (54). This suggests that *in vivo* EcDsbC could be involved in responding to copper stress. Excess copper catalyses disulfide bond formation, including the formation of disulfide bonded oligomers, in the bacterial periplasm (54). EcDsbC is thought to contribute to copper resistance by resolving these copper-catalysed non-native disulfide bonds that would otherwise disrupt the folding and function of many proteins (54). It has been shown that under oxidative stress EcDsbC protects the single cysteine of AraF (an arabinose binding protein), by reducing disulfide-bonded AraF dimers (55).

1.1.5 *E. coli* DsbD

EcDsbC can only function as a disulfide isomerase if its catalytic cysteines are in the reduced form. EcDsbD is responsible for reducing the EcDsbC cysteines and transporting electrons from the cytoplasm to the periplasm to maintain other periplasmic proteins in the reduced state (56). Like EcDsbB, EcDsbD is an inner membrane protein, though its architecture is very different. EcDsbD consists of two periplasmic domains, one at the N-terminus (EcDsbDα) and another at the C-terminus (EcDsbDγ), which are connected by a transmembrane domain (EcDsbDβ) consisting of eight predicted transmembrane helices (57) (Figure 1.8). Each of these domains has two redox active cysteines that are required for the electron transporting activity of EcDsbD (57,58). The expression of all three domains individually is able to complement the loss of the entire EcDsbD protein, suggesting that the domains do not need to be physically linked for its electron transporting function (59,60). The structures of both soluble domains have been solved showing that EcDsbDα consists of an immunoglobulin-like fold (47,61) whereas EcDsbDγ adopts a thioredoxin-like fold (Figure 1.8) (62,63). In both of these domains one catalytic cysteine is more solvent accessible than the other (61,63). These solvent accessible cysteines interact directly with other domains/substrates whereas the shielded cysteines resolve the intra/interdisulfide bonds (44,64). No structures of the full-length protein or of the membrane-embedded domain, EcDsbDβ, have been solved. The catalytic cysteines of EcDsbDβ, Cys163 and Cys285, are located on transmembrane helices 1 and 4, respectively (57). Solvent accessibility studies of EcDsbD have suggested that these helices form an hourglass-shaped pore with the catalytic cysteines in the centre, allowing them to be accessed from both sides of the membrane (inset, Figure 1.8) (65,66).

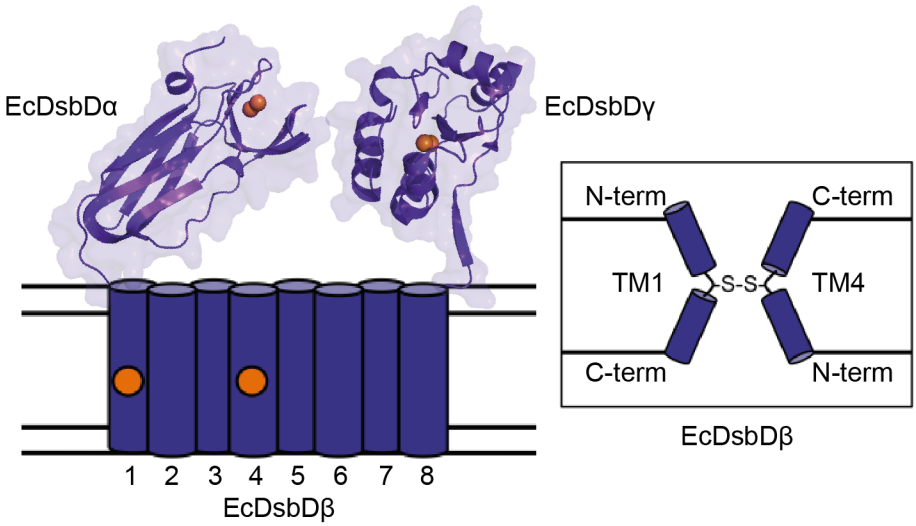


Figure 1.8: Structure of EcDsbD. *The crystal structures of EcDsbDα (PDB ID: 1JPE) and EcDsbDγ (PDB ID: 1UC7) are shown with the catalytic cysteines represented as orange*

spheres. The predicted topology of EcDsbDβ is shown, with the transmembrane helices numbered. Inset shows the predicted hourglass-shaped structure of transmembrane (TM) helices 1 and 4.

EcDsbDα is the domain that is responsible for directly interacting with and reducing EcDsbC. The structure of the EcDsbC-EcDsbDα interaction complex has been solved showing that EcDsbDα interacts with both EcDsbC protomers within the EcDsbC homodimer (Figure 1.9A) (44,47). The interaction between EcDsbC and EcDsbDα results in a disulfide exchange and, like the EcDsbA-substrate interaction (44), this exchange is predicted to proceed via two S_N2 mechanisms (Figure 1.9B). Initially, the thiolate anion of Cys109 on EcDsbDα attacks the EcDsbC disulfide bond, forming an intermolecular disulfide bond between Cys109 (EcDsbDα) and Cys98 (EcDsbC) (61). Cys103 of EcDsbDα then resolves this bond to release reduced EcDsbC and oxidised EcDsbDα (61). EcDsbDα has other interaction partners including EcDsbG, another homodimeric thioredoxin-fold containing protein, which functions to stop the oxidation of single cysteine residues to sulfinic and sulfonic acids (67,68), and EcCcmG (also known as EcDsbE), which is involved in the maturation of cytochrome C (56,69). When *dsbD* is deleted from *E. coli*, these proteins remain oxidised and are unable to perform their function (6,56,68).

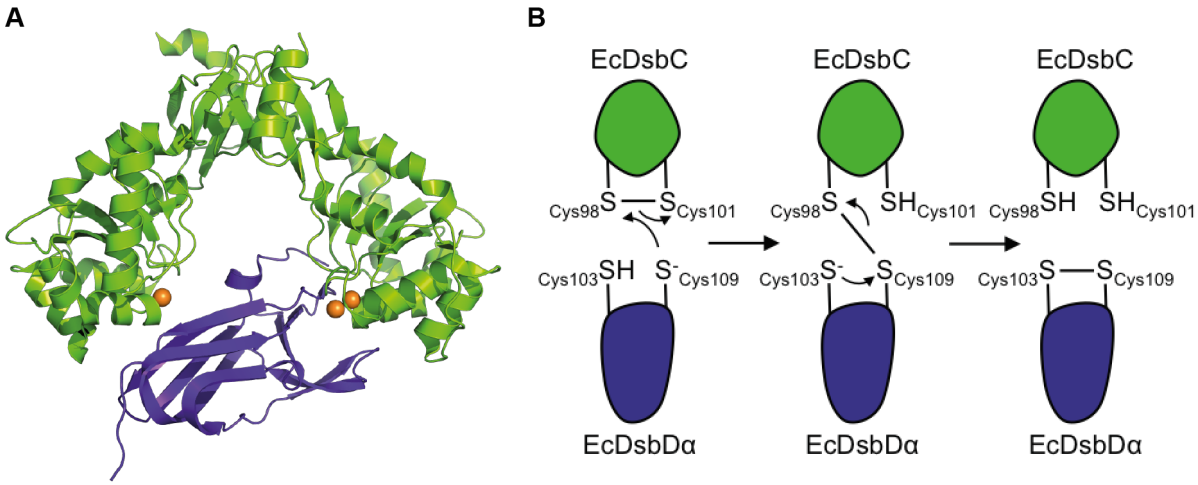


Figure 1.9: The interaction between EcDsbC and EcDsbDα. A. *Crystal structure of the EcDsbC-EcDsbDα complex. EcDsbC is shown in green and EcDsbDα is coloured purple. The catalytic cysteines are shown as orange spheres.* ***B.*** *Proposed mechanism of the disulfide exchange reaction between EcDsbC and EcDsbDα.*

EcDsbDα gains its reducing power from a disulfide exchange cascade that originates in the cytoplasm with the reduction of thioredoxin by NADPH-dependent thioredoxin

reductase (Figure 1.10). The electrons are then passed from thioredoxin to EcDsbDβ, EcDsbDβ to EcDsbDγ and finally from EcDsbDγ to EcDsbDα (59). The mechanism for these disulfide exchanges is predicted to be similar to the EcDsbDα-EcDsbC interaction (59).

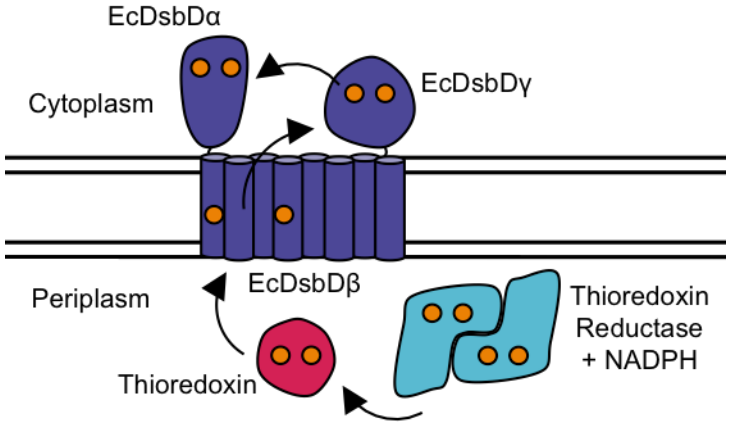


Figure 1.10: EcDsbD disulfide exchange cascade. *EcDsbD transports electrons from the cytoplasm to the periplasm via a disulfide exchange cascade. The curved arrows represent the disulfide exchange reactions. The catalytic cysteines are shown in orange.*

1.1.6 Other *E. coli* Dsb proteins

Other Dsb proteins found in *E. coli* include DsbG and CcmG (also known as DsbE). As mentioned in the previous section, EcDsbD maintains both of these proteins in the reduced state (68). EcDsbG is a dimeric thioredoxin-fold protein with structural features similar to those of EcDsbC (70), despite only 26% sequence identity between the two proteins (Figure 1.11A) (67). The main difference is that EcDsbG has a longer α-helix connecting the catalytic and dimerisation domain, meaning that the cleft between the catalytic domains is significantly larger (~20 Å wider and ~6 Å deeper) (70). The cleft of EcDsbG also contains charged residues suggesting that the protein may interact with folded, or at least partially folded substrates (70). As overexpression of *dsbG* is able to partially complement *dsbC* deletion in *E. coli*, it was originally suggested that EcDsbG functions as an isomerase with a much narrower range of substrates (68). This was supported by the findings that EcDsbG is expressed at much lower levels than EcDsbC *in vivo* (68) and that EcDsbG is not involved in the folding of RNase I (17) or able to function in the insulin reduction assay (68). More recently it was shown that EcDsbG functions as a chaperone (71,72) and that it protects single cysteine residues from oxidation to sulfinic and sulfonic acids (67). This latter study identified three EcDsbG substrates, YbiS, ErfK and YnhG (67), which are all L,D-transpeptidases involved in cell wall synthesis and require a single, reduced cysteine residue for their catalytic function (73).

CcmG stands for cytochrome c maturation protein G and is one of eight Ccm proteins encoded in *E. coli* (74). This protein is also called DsbE because of its similarities to the disulfide bond forming proteins and its interaction with EcDsbD. EcCcmG is a reductase but is considered to have a specialised function interacting with EcCcmH (75), to pass electrons onto apocytochrome c, reducing its CXXCH motif and allowing heme attachment (76,77). EcCcmG consists of a periplasmic C-terminal domain, which is tethered to the inner membrane by a single N-terminal transmembrane α-helix (75). The structure of the C-terminal domain (residues 19-185) has been solved revealing that it contains a thioredoxin fold (Figure 1.11B) (78).

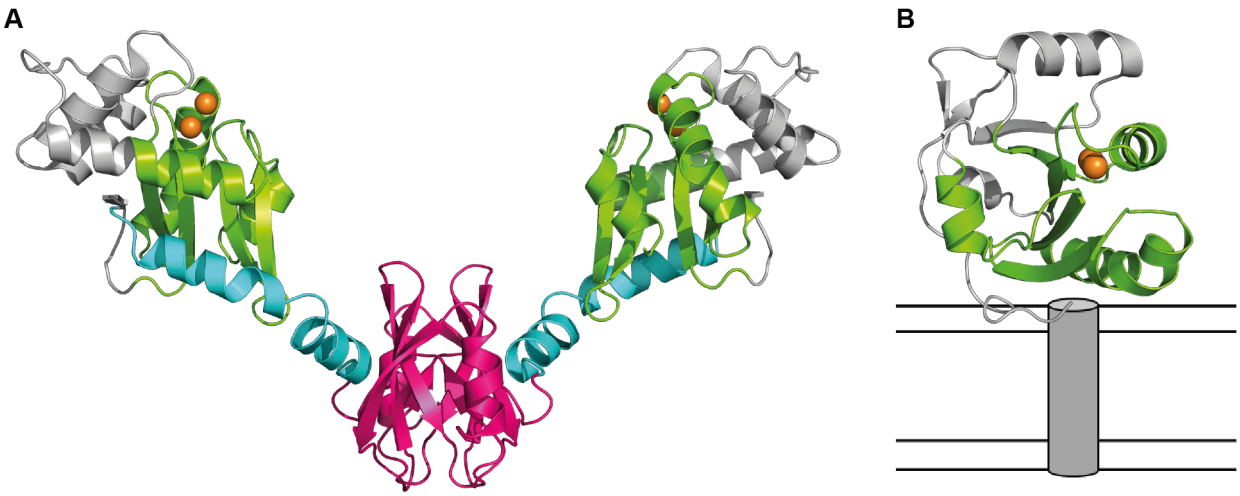


Figure 1.11: Structure of EcDsbG and EcCcmG. A *Structure of EcDsbG (PDB ID: 1V57) with the dimerisation domain coloured magenta, the linker helix coloured cyan and the thioredoxin fold shown in green with the catalytic cysteines shown as orange spheres. Other elements of the catalytic domain are coloured grey.* ***B.*** *Structure of the C-terminal periplasmic domain of EcCcmG (PDB ID: 2B1K). The thioredoxin fold is shown in green and the catalytic cysteines are represented as orange spheres. Other structural elements are coloured grey and the single transmembrane helix is represented as a cylinder to show the topology of the full-length protein.*

1.1.7 Conservation and diversity in the Dsb protein family

Most of the published work on the Dsb family has been performed in the non-pathogenic *E. coli* K12 strain. However, research on homologues of these proteins has suggested this organism may not be representative of disulfide bond formation in all bacteria (79). The Dsb oxidation and isomerisation pathways are generally conserved in gamma- and beta-proteobacteria, but are more diverse in alpha-, delta- and epsilon-proteobacteria and in

Gram-positive bacteria (16,79,80). No DsbC/D-like disulfide isomerisation pathway has been identified in Gram-positive bacteria (79). *Staphylococcus aureus* does not have a detectable DsbB homologue and may instead rely on small molecules to reoxidise DsbA (81-83). Some bacteria have multiple DsbA/B/C/D-like or sets of Dsb-like proteins, which can each have specialised redox functions *in vivo*. For example uropathogenic *E. coli* and *Salmonella enterica* serovar Typhimurium, in addition to the prototypical system, have another set of DsbA/B-like proteins, DsbI/L, that are encoded in an operon and are necessary for the oxidative folding of arylsulfate-sulfotransferase (AssT) (84-87). Another group of accessory Dsb-like proteins are the suppressor of copper sensitivity proteins that contribute to the bacterial virulence trait of copper resistance, and are the focus of this thesis.

1.2 Copper in infection and immunity

Copper (Cu) is an essential trace metal, needed for cellular processes in all aerobic forms of life. Its ability to cycle between two oxidation states (Cu (I) and Cu (II)) makes it a useful cofactor in many enzymes. The concentration of cellular copper must be tightly regulated as the redox activity of this metal also makes it toxic. Copper-mediated toxicity is utilised by the innate immune system to overcome infection (88). Animals deprived of copper in their diets are more susceptible to infection (89-91), whereas increasing dietary Cu intake (while maintaining non-toxic levels) has the opposite effect (92,93). Work has suggested that phagocytes may use Cu ions to kill pathogens that have been engulfed (94-96). In response to bacterial infection higher levels of copper can be found in the extracellular environment and the site of infection. For example, patients with urinary tract infections have increased levels of copper in their urine compared to healthy controls (92,97) and guinea pigs infected with *Mycobacterium tuberculosis* have increased copper concentrations in inflammatory lesions (93).

1.2.1 The antibacterial nature of copper

Copper is toxic to bacterial cells for three main reasons, two of which rely on the redox properties of copper. Firstly, redox cycling between free Cu(I) and Cu(II) can occur in bacteria causing unwanted redox reactions (88), including the formation of non-native disulfide bonds (54,98). Interaction of Cu(I) with hydrogen peroxide or superoxide produced by the host also results in the production of hydroxyl radicals and causes severe oxidative stress to the bacteria (94). The third, redox-independent, effect is that Cu(I) has a greater affinity for thiols than other metals (e.g. iron) and therefore outcompetes them for coordination sites in metalloproteins (88,94,99). This disrupts complexes (e.g. Fe-S clusters) that are required for essential bacterial cellular processes (100).

1.2.2 Bacterial response to copper stress

Bacteria deploy various strategies to evade the toxic effects of copper and these vary significantly across different bacterial species, with some having redundant systems in place (88). Four main features of the bacterial response to copper stress are copper sensors (A), oxidases (B), binding proteins (C) and transporters (D) (Figure 1.12) (88,99). Copper sensors, such as CueR in *E. coli* and *S.* Typhimurium, are activated by an increased concentration of copper in the bacteria and often act as transcription factors that turn on genes encoding other copper response proteins like copper binders and transporters (94,101-103). CueO is an example of a copper oxidase, which functions in the

periplasm of bacteria to oxidise Cu(I) to Cu(II), which is less toxic (104,105). Those proteins that bind excess copper, such as CueP in *S.* Typhimurium (106), either sequester it or chaperone it to other proteins, such as copper sensors or transporters/efflux systems that remove it from the cell (99). A key feature of many bacterial copper response systems is the presence of CopA, a Cu(I)-exporting ATPase that transports Cu(I) from the cytoplasm to the periplasm (88,107,108). Other less conserved copper exporters, such as *E. coli* CusCFBA, pump copper out of the periplasmic space (109,110). Some bacteria also have additional machinery that contributes to copper resistance in other ways; this machinery includes DsbC in *E. coli* (as mentioned previously) and the understudied suppressor of copper sensitivity proteins.

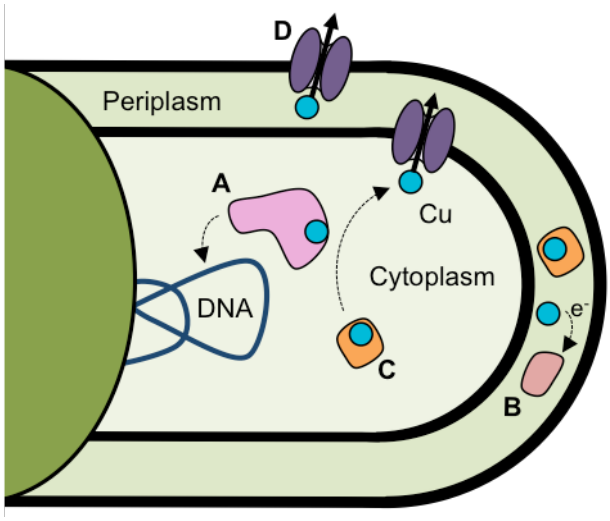


Figure 1.12: Copper resistance mechanisms in bacteria. A. *Metal sensors detect copper ions in the cytoplasm and often act as transcription factors.* ***B****. Some periplasmic proteins oxidise Cu (I) to Cu (II).* ***C.*** *Copper binders sequester copper or chaperone it to other copper resistance proteins.* ***D.*** *Copper transporters can remove copper ions from the cytoplasm and periplasm.*

1.3 Suppressor of copper sensitivity proteins

Studies of the suppressor of copper sensitivity (Scs) proteins have shown that they are involved in the bacterial response to copper and oxidative stress, but their specific cellular role is unknown. The Scs proteins have similarities to the Dsb proteins but are not present in *E. coli*, and are encoded at a single locus (111). This locus has four protein encoding genes and was first identified in *S.* Typhimurium, when it was shown to increase the copper tolerance of copper sensitive *E. coli* mutants (111). These copper sensitive *E. coli* strains had mutations in either *cutA* (encoding DsbD), *cutC* (encoding a copper homeostasis protein), *cutE* (encoding the apolipoprotein N-acyltransferase Lnt), *cutF* (encoding an outer membrane lipoprotein NlpE) or *lgt* (encoding a lipoprotein diacylglyceryl transferase) (111). Recently the transcription of the *scs* genes in *S.* Typhimurium, was shown to be upregulated in the presence of copper and dependent on the transcription factor CpxR (112). CpxR, in concert with CpxA, senses misfolded periplasmic proteins and allows the transcription of the Cu(II) binding protein CueP and other proteins under copper stress (113-116). This study also found that a promoter upstream of *scsA*, drives the transcription of all *scs* genes. Deletion of the *scs* locus in *S.* Typhimiurium results in reduced growth of the bacterial pathogen under copper stress (112,117). However, one study found that the Δ*scs* strain had increased invasiveness *in vitro* and in a mouse model of infection (117). Overall, animals infected with the Δ*scs* strain had a similar death rate to those infected with wild-type *S.* Typhimurium (117).

The four proteins encoded in the *S.* Typhimurium *scs* locus, ScsA, ScsB, ScsC and ScsD, all contain cysteine residues predicted to form redox-active pairs and three harbour predicted thioredoxin folds (111). The Scs proteins are encoded in a number of pathogenic bacteria (112,118), though little is known about these proteins. Most of the published work has focused on the Scs systems in *S.* Typhimurium and, to a lesser extent, *Caulobacter crescentus*.

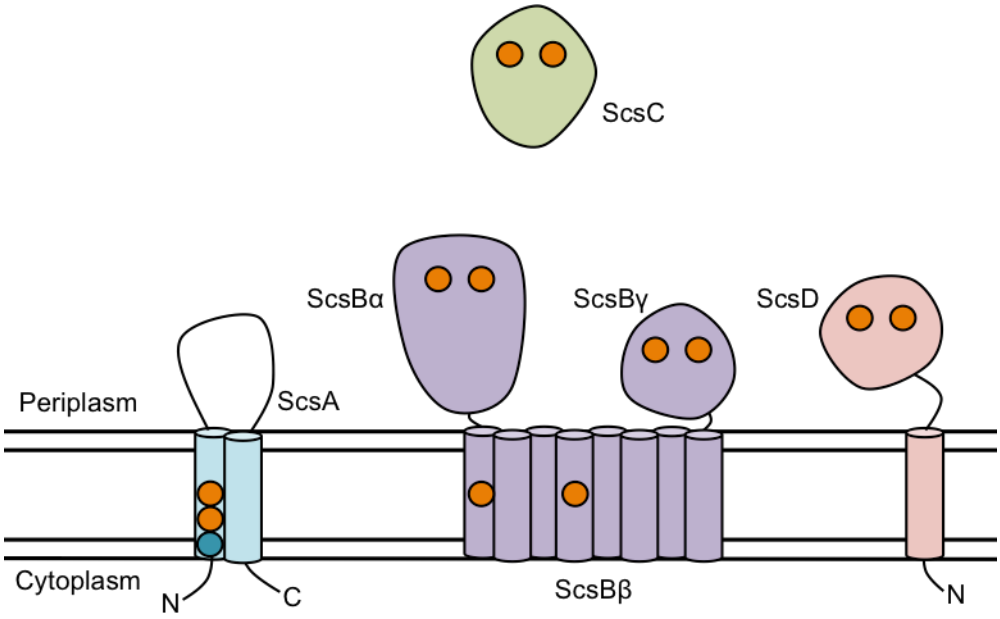


Figure 1.13: Predicted architecture of the Scs proteins. *In each protein the predicted catalytic cysteines are represented as orange circles and the cyan circle in ScsA represents a predicted copper-binding motif.*

1.3.1 ScsA

S. Typhimurium ScsA (StScsA) has no overall homology to any structurally characterised protein and was originally predicted to be a membrane protein with two transmembrane helices (Figure 1.13, residues 10-30 and 71-91) (111). The protein also has a CXXC motif and putative copper-binding motifs (1-MAKQQRM-7 and 20-MVCTAQRMAGLHALQM-35) but these reside in the first predicted transmembrane helix (111). More recent analysis suggests that the protein has a lipoprotein signalling sequence and if this were cleaved *in vivo*, it would remove both the CXXC motif and copper-binding motif (117). No biochemical or structural analyses have been performed on any ScsA protein to date, however, genetic analysis has provided some insight into its biological function. Deletion of *scsA* (Δ*scsA*) in *S*. Typhimurium does not increase the sensitivity of the bacteria to copper, revealing that the encoded protein does not play a role in the response to copper stress (112,117). Two studies have found that *scsA* deletion mutants are sensitive to H_2O_2 (which generates reactive oxygen species) suggesting that ScsA plays a role in the response to oxidative stress (112,117). When exposed to 6 mM H_2O_2 the Δ*scsABCD S.* Typhimurium strain displays the same phenotype as the Δ*scsA* strain (112). This deletion mutant also has increased protein carbonylation when exposed to H_2O_2 (117), a key marker of oxidative stress. Another study suggests that *scsA* is involved in cortisol-induced proliferation of *S.*

Typhimurium in host macrophages and showed that deletion of *scsA* results in upregulation of *scsBCD* (119).

1.3.2 ScsB

The second open reading frame in the *scs* locus encodes the predicted inner transmembrane protein; ScsB (111). Phylogenetic analysis identified ScsB as a new class of DsbD-like protein (118). The predicted architecture of ScsB is very similar to that of EcDsbD. Like EcDsbD, ScsB is predicted to have N-terminal (α) and C-terminal (γ) periplasmic domains that are separated by a transmembrane domain (β) (Figure 1.13) (118). ScsBβ is predicted to consist of 6 or 8 transmembrane helices depending on the organism of origin. ScsBγ is predicted to have a thioredoxin fold and all domains contain two putative redox active cysteines, just like DsbD. This suggests that ScsB has a similar function to DsbD in transporting electrons from the cytoplasm to the periplasm (59). The key difference between ScsB and DsbD is the α-domain (118). In DsbD, it is this domain that directly interacts with substrates (59). Curiously, ScsBα is much larger than DsbDα, with approximately 90 more residues, suggesting that they have different structures and act on different substrates (118). Three substrates of *C. crescentus* ScsB (CcScsB) have been identified. These are the peroxiredoxin-like1 protein (PrxL), a thioredoxin-like protein TlpA and ScsC, all of which are reduced by CcScsB (118). The function of PrxL is unknown. TlpA reduces peroxiredoxin (PprX), which in turn reduces both hydrogen peroxide and organic peroxides (118). This suggests that CcScsB may have a role in facilitating protection from oxidative stress caused by peroxides in the bacterial periplasm.

The deletion of *scsB* in *S.* Typhimurium results in a slightly extended lag phase in the growth of the bacteria when exposed to H_2O_2, suggesting that ScsB is somehow involved in overcoming oxidative stress in this organism too (112). Deletion mutants of *S.* Typhimurium suggest that ScsD as well as ScsC may be a substrate of ScsB. Deletion of *scsB* alone in *S.* Typhimurium results in the same level of copper sensitivity as the *scsABCD* deletion mutant, whereas the individual Δ*scsC* and Δ*scsD* mutants show lower levels of sensitivity (112). The Δ*scsCD* mutant, however, has a similar level of copper sensitivity to Δ*scsB*, suggesting that the function of ScsC and ScsD under copper stress is dependent on ScsB (112). Prior to the work described in this thesis, no structural studies of ScsB had been published.

1.3.3 ScsC

S. Typhimurium ScsC (StScsC) was originally predicted to be an outer membrane protein (111), but more recent work showed that it is a globular periplasmic protein with oxidoreductase activity (120). StScsC is monomeric, with a DsbA-like structure (Figure 1.14A). The StScsC structure consists of a thioredoxin fold, harbouring the CXXC catalytic motif on α-helix 1 and an inserted helical domain (120). Notably, the StScsC structure is considerably different from *S.* Typhimurium DsbA and EcDsbA. Key differences in the StScsC structure include the presence of a short helical domain at the N-terminus, a change in position of β1, which interacts with β3 rather than β5, and the presence of a loop in place of the fourth α-helix of the helical insertion in EcDsbA (120). These features are generally more common in disulfide isomerases like EcDsbG and EcDsbC than oxidases (39,70), though the overall structure is reported to be similar to a DsbA-like protein, BdbD, from the Gram-positive bacteria *Bacillus subtilis* (120). The catalytic motif of StScsC is most similar to EcDsbG, as both proteins have a proline and a tyrosine separating the two catalytic cysteines. This explains why the redox potential of StScsC (-132 mV) is lower than that of EcDsbA (-122 mV), as a tyrosine residue in the second X position of the CXXC motif is thought to be less stabilising for the reactive thiolate anion than the histidine in EcDsbA (70, 120). Another similarity between StScsC and EcDsbG is that a lysine residue may be involved in stabilising the thiolate anion form of the solvent accessible catalytic cysteine (Figure 1.14) (120).

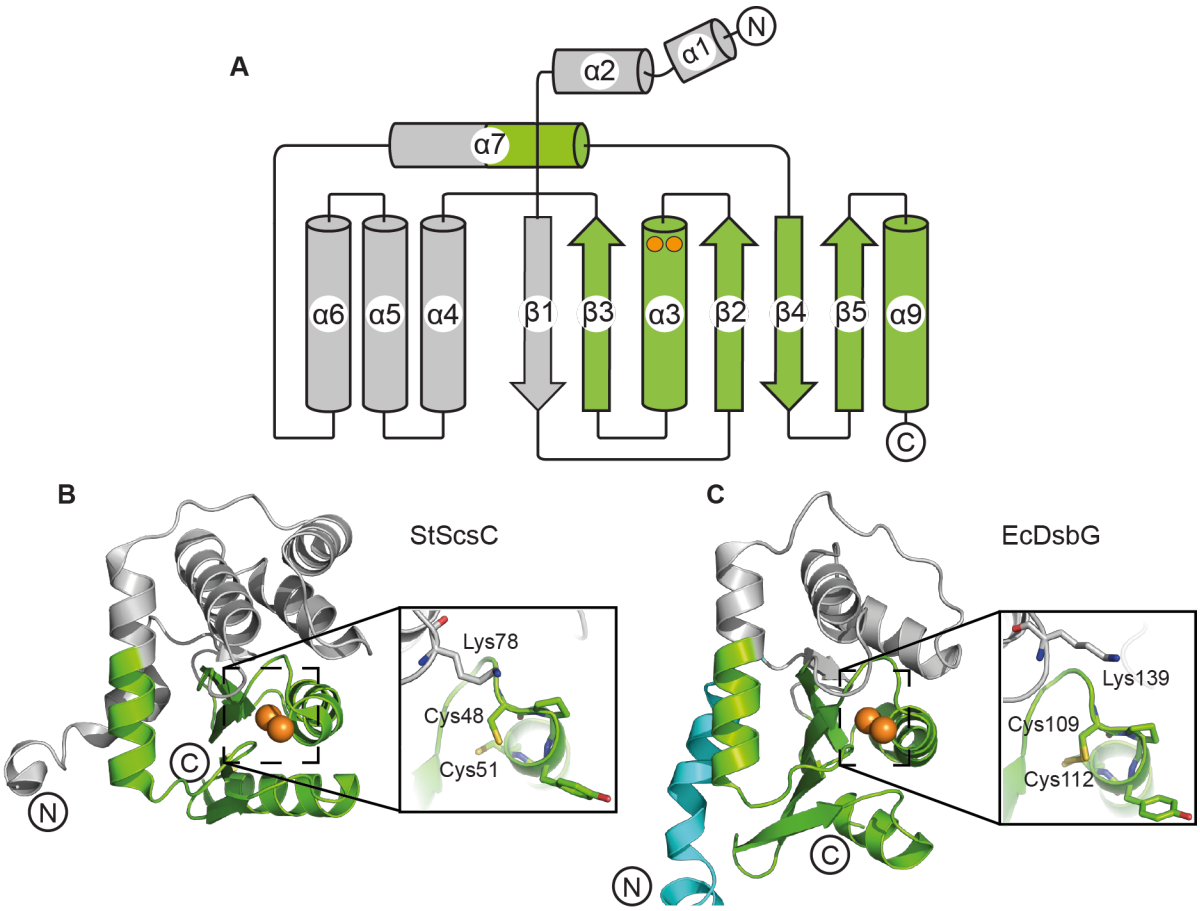


Figure 1.14: Structure of StScsC. A. *Topology diagram of StScsC.* ***B.*** *Structure of StScsC (PDB ID: 4GXZ), with inset showing the catalytic site.* ***C.*** *Structure of a catalytic domain of EcDsbG (PDB ID: 1V57), with inset showing the catalytic site, for comparison to StScsC. In all panels, the thioredoxin fold is coloured green, other structural features of the catalytic domain are coloured grey and the catalytic cysteines are represented as orange circles or spheres. The N- and C-termini are labelled in each panel.*

StScsC has biochemical properties that are typical of Dsb oxidoreductases, including a redox potential of -132 mV and a pKa of 3.4 for the solvent accessible catalytic cysteine (120), closely resembling those of EcDsbG (redox potential -129 mV (121) and pKa 3.5 (122)). Though, unlike EcDsbG, StScsC can catalyse the reduction of insulin, suggesting that it has a broader substrate range (68,120). StScsC is predominantly in the reduced state *in vivo* however it does not function as an isomerase in the scrambled RNase A refolding assay (120). This may be a result of the protein being monomeric, as bacterial disulfide isomerases typically are dimeric (123). Loss of *scsC* results in reduced growth of *S.* Typhimurium in the presence of copper (112,117,120) and, as was shown for *scsB*, deletion of *scsC* also results in slightly increased sensitivity to H_2O_2 (112).

ScsC from *C. crescentus* (CcScsC) has also been characterised, revealing that it is structurally and functionally different to StScsC (118). Though the structure of CcScsC has not been solved, analysis of the protein using size exclusion chromatography suggests that it forms a dimer in solution (118). Secondary structure prediction revealed that the C-terminal domain contains a thioredoxin fold and the N-terminus forms a long alpha helix, which is predicted to facilitate dimerisation (118). This N-terminal alpha helix is shorter in StScsC, which may explain why it is monomeric. CcScsC also functions as a disulfide isomerase and is maintained in the reduced state *in vivo* by CcScsB (118). This suggests that there is a DsbC-DsbD like interaction between ScsC and ScsB in *C. crescentus* though this interaction has not been confirmed in organisms with a monomeric ScsC.

Bioinformatic analysis of ScsC sequences shows great diversity in homologues, suggesting that it may have diverse structures and be involved in a variety of redox reactions across different groups of bacteria (118).

1.3.4 ScsD

Very little work has been done towards understanding the role of ScsD. However, ScsD does have homology to a number of other proteins, which give some insight into its potential function (111). ScsD is predicted to consist of a periplasmic thioredoxin-like domain anchored to the membrane by a single N-terminal transmembrane helix (Figure 1.13) (111). This architecture is very similar to that of CcmG, which is a thiol reducant and plays a role in cytochrome c biogenesis (69,75,124). *S.* Typhimurium ScsD is homologous to thioredoxin-like proteins involved in cytochrome biogenesis in other bacteria including TlpA from *Bradyrhizobium japonicum* and various protein disulfide isomerases (111). This information suggests that ScsD is involved in the reduction of proteins and potentially involved in cytochrome biogenesis in the bacterial periplasm. Given that TlpA is a substrate of ScsB in *C. crescentus* (118) and EcCcmG is reduced by EcDsbD (125), it is also predicted that ScsD interacts with ScsB. Finally, gene deletion studies in *S.* Typhimurium show that ScsD plays a role in resistance to copper and to a lesser extent, oxidative stress (112,117).

1.4 Significance

The Scs proteins are a family of understudied Dsb-like proteins that have been shown to play a role in the bacterial response to copper and oxidative stress. This thesis describes the first studies that characterise the Scs proteins from the bacterial pathogen, *Proteus mirabilis*.

P. mirabilis is a Gram-negative enteric bacterium that causes urinary tract infections (UTIs), particularly in people with catheters (126,127). Infection can lead to bacteraemia, infection of the bladder (cystitis) or kidneys (pyelonephritis) (128) and often, urinary stones (129), which severely complicate UTIs (126). *P. mirabilis* infection of the urinary tract is aided by a number of key virulence factors (127). A characteristic feature of *P. mirabilis* is its ability to differentiate into swarmer cells; highly flagellated cells that enable the bacteria to be extremely motile and move throughout the host. The bacterium also uses a number of adhesion factors to colonise the urinary tract (130) and urease, which catalyses the formation of urinary stones (129). Antibiotics such as quinolones and cephalosporins can be used to treat *P. mirabilis* UTIs; however, multi-drug resistant strains have been identified (131,132) and pose a threat to successful treatment of these infections. DsbA is essential for the correct formation of virulence factors in *P. mirabilis*, such as fimbriae for adhesion (133). *P. mirabilis* DsbA has been characterised (134) and this organism also has a DsbB and three predicted DsbC-like proteins. BLAST searches revealed that *P. mirabilis* does not have a DsbD protein; however, it does encode all four suppressor of copper sensitivity proteins, including the DsbD-like ScsB and DsbA-like ScsC. These proteins may play a key role in the response of *P. mirabilis* to copper stress, especially during infection of the urinary tract when host urine copper levels are elevated (97).

Overall, the suppressor of copper sensitivity proteins from *P. mirabilis* are of interest because of their potential role in the virulence traits of the pathogen, including its response to copper stress. The findings outlined in the subsequent chapters of this thesis have deepened our understanding of the Scs proteins, which may prove to be useful in future antibacterial, biotechnology and/or bioremediation strategies.

1.5 Aims

Broadly, this thesis focuses on the biochemical and structural characterisation of the ScsC and ScsB proteins from *Proteus mirabilis*.

The specific aims of this thesis are:

1. To characterise the ScsC protein from *P. mirabilis* (PmScsC)
2. To structurally characterise the N-terminal domain of *P. mirabilis* ScsB (PmScsBα)
3. To characterise the interaction between *P. mirabilis* ScsC and ScsBα
4. To develop an expression and purification protocol for the production of the transmembrane protein *P. mirabilis* ScsB

Chapter 2 and Chapter 3 describe the work performed towards Aim 1, Chapter 4 describes the work done to achieve Aim 2 and Aim 3 and Chapter 5 describes the work towards Aim 4 (Figure 1.15).

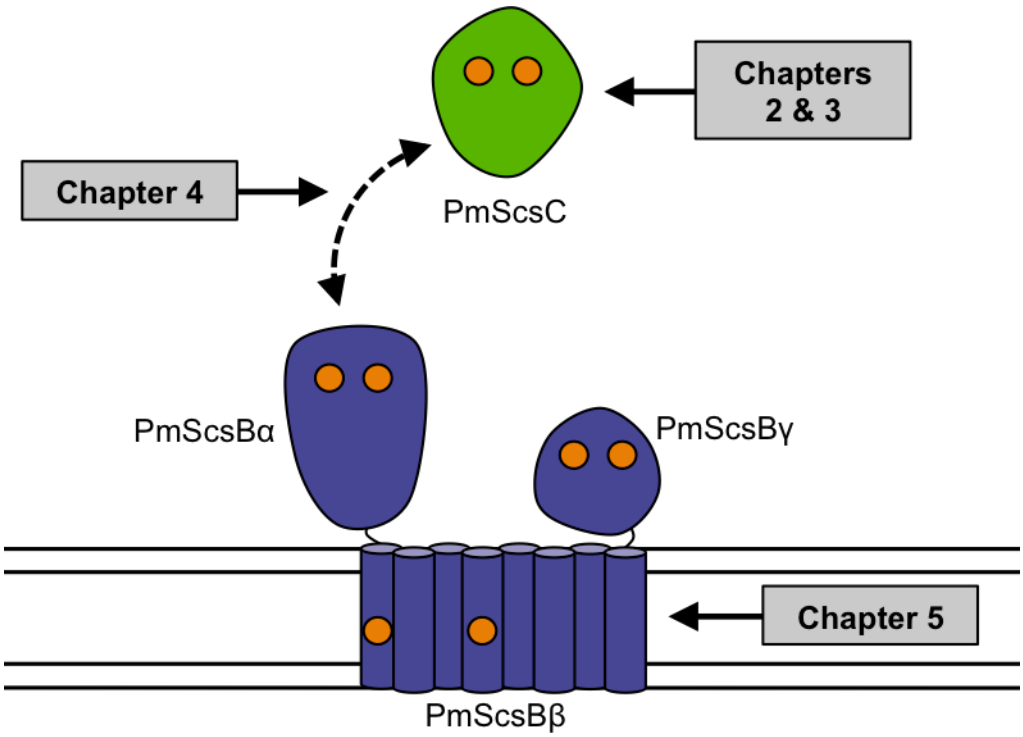


Figure 1.15: Diagram depicting the focus of subsequent chapters in this thesis.

Chapter 2

2. Characterisation of *Proteus mirabilis* ScsC

2.1 Chapter foreword

This chapter is a reproduction of a research article that was published in the journal *Nature Communications*. This publication describes the functional and structural characterisation of PmScsC, revealing that it is a trimeric, shape-shifting disulfide isomerase. Changes have been made to the paper to conform to the thesis submission requirements; these include the removal of author contributions (now listed in the thesis preliminary pages), the overall formatting and the figure, table and reference numbering.

2.1.1 Reference

Furlong, E. J., Lo, A. W., Kurth, F., Premkumar, L., Totsika, M., Achard, M. E. S., Halili, M. A., Heras, B., Whitten, A. E., Choudhury, H. G., Schembri, M. A., and Martin, J. L. (2017) A shape-shifting redox foldase contributes to *Proteus mirabilis* copper resistance. *Nature Communications* **8**, 16065; doi: 10.1038/ncomms16065.

2.1.2 Contribution

I substantially contributed to the analysis and interpretation of the research data outlined in this publication. I produced the crystals that lead to the transitional and extended PmScsC structures, collected diffraction data, then processed the data and solved and refined these crystal structures. My analysis of all three structures forms a major part of the publication. In addition to this, I designed the PmScsC RHP mutant, created the PmScsC RHP expression construct, produced the protein used in experiments and biochemically characterized the *in vitro* activity of PmScsC RHP and interpreted these results. I also had a substantial role in drafting the manuscript, designing and creating the figures and formatting the manuscript for publication.

A shape-shifting redox foldase contributes to *Proteus mirabilis* copper resistance

Emily J. Furlong[#], Alvin W. Lo[¶*], Fabian Kurth[#], Lakshmanane Premkumar[#2], Makrina Totsika[¶*], Maud E. S. Achard[¶*], Maria A. Halili[#], Begoña Heras[§], Andrew E. Whitten[#1], Hassanul G. Choudhury[#], Mark A. Schembri[¶*1], and Jennifer L. Martin[#%1]

[#]Institute for Molecular Bioscience, University of Queensland, St Lucia, QLD 4072, Australia. [¶]School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, QLD 4072, Australia. *Australian Infectious Diseases Research Centre, University of Queensland, St Lucia QLD 4072, Australia. [§]La Trobe Institute for Molecular Science, La Trobe University, Bundoora, VIC 3068, Australia. [%]Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD 4111 Australia.

Currently: FK Bristol-Myers Squibb, Arnulfstraße 29, 80636 Munich, Germany; LP Department of Microbiology and Immunology, School of Medicine, University of North Carolina, Chapel Hill, NC 27514 U.S.A.; MT Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Queensland University of Technology, Kelvin Grove, QLD 4059, Australia; MESA, School of Human Movement and Nutrition Sciences, University of Queensland, St Lucia, QLD 4072, Australia; AEW Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW 2234, Australia. HGC Cello Health Consulting, Farnham Surrey GU9 7DN, U.K.

[1]To whom correspondence should be addressed:
(jlm@griffith.edu.au; m.schembri@uq.edu.au; awh@ansto.gov.au)

Publication timeline
Received for publication: 7 October 2016
Accepted: 24 May 2017
Published: 19 July 2017

2.2 Abstract

Copper resistance is a key virulence trait of the uropathogen *Proteus mirabilis*. Here we show that *P. mirabilis* ScsC (PmScsC) contributes to this defense mechanism by enabling swarming in the presence of copper. We also demonstrate that PmScsC is a thioredoxin-like disulfide isomerase but, unlike other characterised proteins in this family, it is trimeric. PmScsC trimerisation and active site cysteine are required for wild-type swarming activity in the presence of copper. Moreover, PmScsC exhibits unprecedented motion as a consequence of a shape-shifting motif linking catalytic and trimerisation domains. The linker accesses strand, loop and helical conformations enabling the sampling of an enormous folding landscape by the catalytic domains. Mutation of the shape-shifting motif abolishes disulfide isomerase activity, as does removal of the trimerisation domain, showing that both features are essential to foldase function. More broadly, the shape-shifter peptide has the potential for “plug and play” application in protein engineering.

2.3 Introduction

The *Escherichia coli* periplasmic Dsb protein family is the best characterised bacterial oxidative folding system. *E coli* DsbA is a monomeric thioredoxin-fold oxidase (7) that introduces disulfide bonds into protein substrates (25) whereas *E. coli* DsbC is a dimeric V-shaped thioredoxin-fold protein disulfide isomerase (39) that proof-reads and shuffles incorrect disulfide bonds (4). *E. coli* DsbG is also a dimeric V-shaped thioredoxin-fold protein disulfide isomerase and cysteine reducing system (70) that protects cysteines of periplasmic proteins from inappropriate oxidation (67).

Here, we characterise a putative DsbA-like protein from *P. mirabilis* and show it plays a key role in virulence by enabling swarming during copper stress. We find that it is not DsbA-like, as it does not catalyse disulfide bond formation. Rather, it is a disulfide isomerase that shuffles incorrect disulfide bonds in proteins. Further, we show that this thioredoxin fold protein is unlike any other characterised to date - it is trimeric - and we demonstrate that its function depends on a shape-shifting motif that could potentially be used as a “plug and play” peptide module.

2.4 Results

Proteus mirabilis UniProt C2LPE2 is a predicted thioredoxin-fold protein, and the purified protein exhibits redox properties characteristic of the thioredoxin fold family. It is highly oxidising (–108 mV), has an acidic active site cysteine (pK_a 3.1) and the oxidised disulfide form of the active site destabilises the protein compared with the reduced form (Supplementary Figure 2.1A-C).

The encoded protein plays an important role in *P. mirabilis* virulence. Copper intoxication is a component of nutritional immunity having a number of detrimental effects on bacterial cells. Copper can react with host-generated hydrogen peroxide to generate free radicals, which damage biological molecules (95); cycling between Cu(I) and Cu(II) can lead to detrimental redox reactions in bacteria; and copper can bind to protein thiols in place of other metal co-factors (88,94). Overcoming these antibacterial effects of copper is a key bacterial defence mechanism. The *P. mirabilis* C2LPE2 is located within a cluster of four predicted **s**uppressor of **c**opper **s**ensitivity (Scs) genes and we proposed that encoded protein (hereafter called PmScsC) would contribute to bacterial copper resistance. Indeed, inactivation of *scsC* in two independent *P. mirabilis* clinical isolates significantly inhibited swarming motility (a key virulence trait of *P. mirabilis*) in the presence of copper (Figure 2.1B, Supplementary Figures 2.2, 2.3 and 2.4). This phenotype could be complemented by introduction of a plasmid encoding wild-type PmScsC, but not a plasmid encoding an inactive PmScsC (Figure 2.1A, B). Although it is annotated DsbA-like, PmScsC is a powerful protein disulfide isomerase (Figure 2.1C) that is able to refold and reactivate the scrambled disulfide form of the model substrate RNase A just as rapidly as the archetypal disulfide isomerase *E. coli* DsbC. It has negligible DsbA-like dithiol oxidase activity (Supplementary Figure 2.1D).

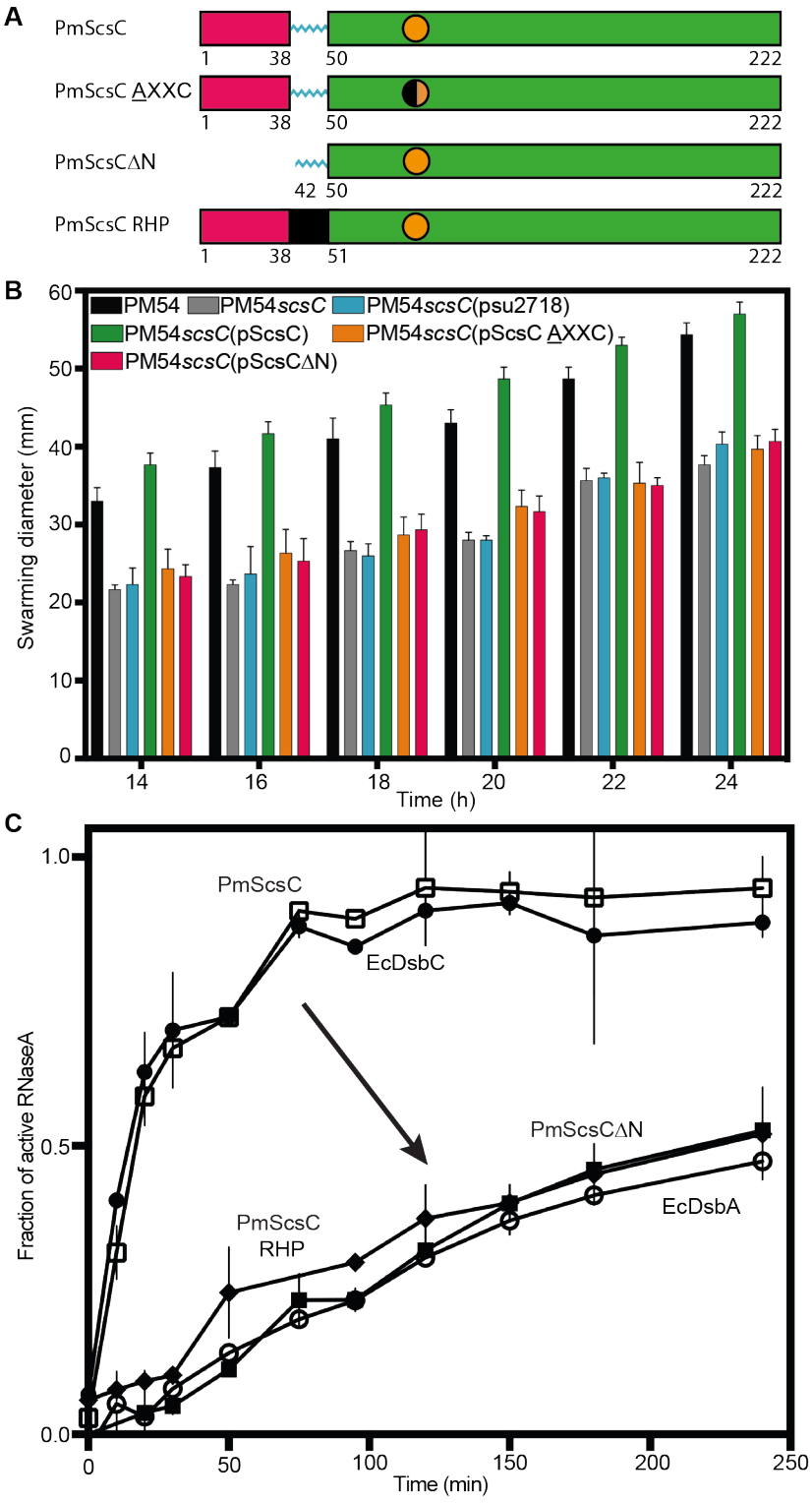


Figure 2.1: PmScsC function. A. *Linear representation of the domain organisation of PmScsC mutants. The trimerisation domain is coloured magenta, the catalytic domain is green, with the CXXC motif represented as an orange circle, and the linker region is shown in cyan. The rigid helical linker is shown in black.* ***B.*** *Swarming motility of wild type P. mirabilis strain PM54 (black), PmScsC deletion mutant PM54scsC (grey), and PM54scsC containing control and complementation plasmids: PM54scsC(pSU2718) (vector control; cyan), PM54scsC(pScsC) (wild-type; green), PM54scsC(pScsC AXXC)*

(active site mutant; orange) and PM54scsC(pScsCΔN) (N-terminal trimerisation domain deletion; magenta) in the presence of 1.5 mM $CuSO_4$ on LB agar plates (significant difference, P < 0.0001 for slope calculated by F-test). Data are shown as the mean±s.d., of a single experiment performed in triplicate; all data is representative of three independent experiments. ***C.*** *Disulfide isomerase activity of PmScsC (hollow squares) is similar to that of EcDsbC (filled circles). Monomeric PmScsΔN (filled squares) and the PmScsC mutant engineered to have a rigid helical peptide (RHP) in place of the flexible peptide (filled diamonds) show negligible activity (equivalent to oxidase EcDsbA (hollow circles)). Data are shown as the mean ± s.d., of two replicate experiments.*

The best-characterised protein disulfide isomerases - EcDsbC and eukaryotic protein disulfide isomerase (PDI) - each have two thioredoxin-fold catalytic domains. DsbC is a dimer, and PDI is a modular 4-domain protein with two catalytically active thioredoxin domains. The presence of two thioredoxin catalytic domains is thought to contribute to highly efficient disulfide shuffling activity (41,43). The amino acid sequence of PmScsC encodes a single thioredoxin fold, and we therefore expected that its disulfide isomerase activity would be a consequence of dimerisation to generate the necessary two catalytic domains. We also expected that the protein would adopt the same dimeric architecture as V-shaped EcDsbC. However, we were wrong on both counts.

Unexpectedly, evidence from chemical cross-linking (Supplementary Figure 2.5), small angle X-ray solution scattering (Figure 2.2A, B) and multi-angle light scattering (Figure 2.2C) all indicated that PmScsC is trimeric. We confirmed that PmScsC is a trimer by determining three independent crystal structures. The three crystal structures – which we refer to as compact, transitional, and extended (Figure 2.3A-C, Table 2.1) - reveal an extraordinary range of motion in this protein. Importantly, the compact and transitional structures have multiple protomers in the asymmetric unit that adopt the same overall structures in each case (8 compact trimers, RMSD ~1.5 Å for 645 Cα; 2 transitional trimers, RMSD 0.5 Å for 653 Cα) so that crystal symmetry does not appear to impact on the observed conformation. Moreover, the compact and transitional conformations were determined from crystals grown under similar conditions (2.85 M sodium malonate pH 5.8, 20 °C with either 0.1 M cobalt or copper added). Nevertheless, comparison of the eight compact and the two transitional trimers in these two crystal structures, reveals an unprecedented level of conformational re-arrangement (RMSDs > 20 Å). The third structure was determined from crystals grown in the presence of 2 mM copper chloride

(32% Jeffamine M-600, 0.1 M HEPES pH 8, 20 °C) and reveals a fully extended trimer with one protomer in the asymmetric unit.

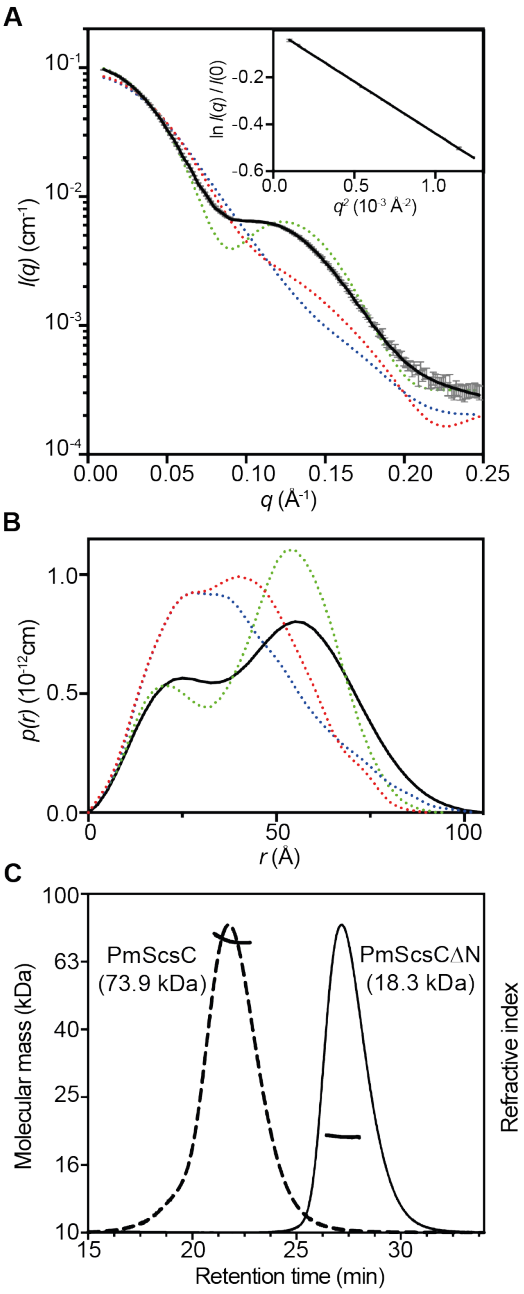


Figure 2.2: SAXS and MALLS of PmScsC. A. *Small-angle X-ray scattering data collected from wild type PmScsC (grey) and the calculated scattering profile of the ensemble model overlayed in black (SASBDB: SASDB94). The predicted scattering profile of each of the crystal structures is also shown (dashed lines: PDB: 4XVW compact, red; PDB: 5IDR transitional, blue; PDB: 5ID4 extended, green). The agreement between the experimental data and the ensemble model is excellent, yielding* χ^2*=1.0 (compared to* χ^2*=863.9 (compact);* χ^2*=1222.0 (transitional);* χ^2*=348.2 (extended)). The Guinier region (inset) of the scattering data is linear, consistent with a monodisperse solution.* ***B.*** *Pair distance distribution function derived from the scattering data, showing the maximum*

dimension of the particles in solution is 105 Å. Also shown is the calculated p(r) for each of the crystal structures (dashed lines: compact, red; transitional, blue; extended, green), showing a maximum dimension of 90, 105 and 100 Å, respectively. The p(r) generated from the extended structure (green) is most similar to the experimentally derived p(r), while the other p(r) curves are markedly different. ***C.*** *MALLS profile of PmScsC and PmScsCΔN. PmScsC eluted faster (21.5 min) than PmScsCΔN (27.2 min): experimentally determined molecular masses are 74 ± 0.3 kDa for PmScsC (theoretical trimer mass 74.3 kDa) and 18.3 ± 0.3 kDa for PmScsCΔN (theoretical monomer mass 18.1 kDa).*

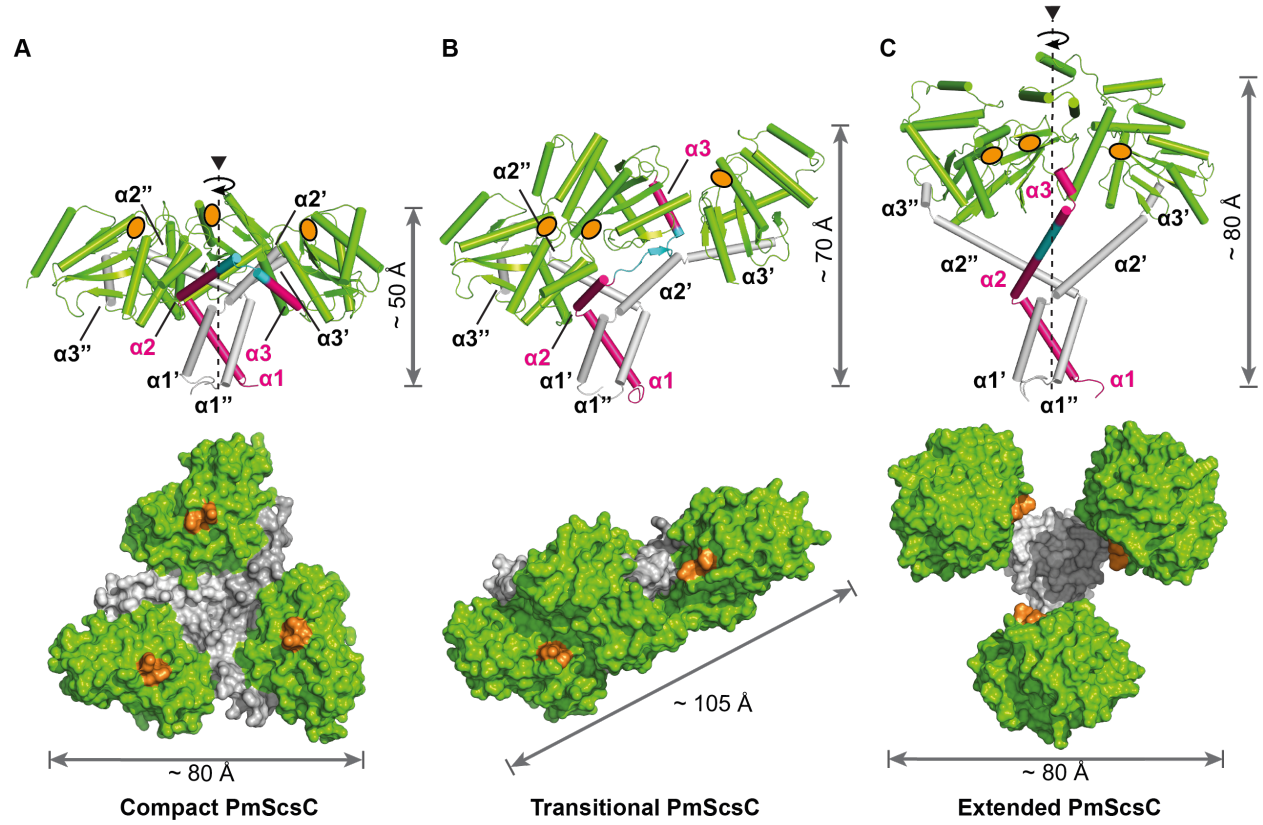


Figure 2.3: PmScsC crystal structures. ***A.*** *Compact (PDB: 4XVW),* ***B.*** *Transitional (PDB: 5IDR) and* ***C.*** *Extended (PDB: 5ID4) crystal structures of PmScsC. Upper panels side view, secondary structure with catalytic domains in green, 11-residue peptide linker in cyan, and trimerisation domains in magenta or white. Height in this orientation is indicated. In each case, one trimerisation domain is shown in magenta for comparison of the conformational changes across the three crystal structures. Lower panels top view, surface representation (catalytic TRX fold domains green, trimerisation domains white), maximum dimension in this orientation is labelled. Active site positions are indicated in orange for each protomer.*

Table 2.1: PmScsC crystal structure statistics.

	Compact (4XVW)	**Transitional (5IDR)**	**Extended (5ID4)**
Data collection			
Space group	P 2_1	I4	H32
Cell dimensions			
a, b, c (Å)	137.5, 163.9, 181.9	193.1, 193.1, 105.8	86.7, 86.7, 330.9
α, *β*, *γ* (º)	90, 90, 90	90, 90, 90	90, 90, 120
Resolution (Å)	91.15 – 2.60 (2.74 – 2.60)	136.51 – 2.56 (2.57-2.56)	110.29 – 2.92 (2.93-2.92)
R_{merge}	0.072 (0.617)	0.083 (0.741)	0.059 (0.625)
$\langle I/\sigma I \rangle$	11.0 (2.0)	14.9 (2.2)	14.2 (2.8)
Completeness (%)	98.6 (95.4)	99.4 (100.0)	99.2 (100.0)
Redundancy	3.8 (3.7)	4.1 (4.1)	4.1 (4.2)
Refinement			
Resolution (Å)	91.15 – 2.60	42.82 - 2.56	40.36 – 2.92
No. reflections	243409	62069	10652
R_{work} / R_{free} (%)	24.8 / 28.2	17.1 / 22.2	25.1 / 26.3
No. atoms			
Protein	40850	10262	1720
Ligand/ion	NA	NA	NA
Water	281	82	0
B factors (Å^2)			
Protein	59.7	50.6	122.2
Ligand/ion	NA	NA	NA
Water	41.5	43.0	NA
RMS deviations			
Bond length (Å)	0.006	0.008	0.010
Bond angles (º)	1.21	1.05	1.17

Single crystals were used to collect each dataset. Values for highest resolution shell are shown in parentheses.

PmScsC also exhibits dynamic conformational flexibility in solution. No single PmScsC crystal structure is consistent with the experimentally determined SAXS scattering curve. However, an ensemble of structures fits the experimental data closely (Figure 2.2A, B) supporting the notion that the trimer is highly dynamic in solution. Moreover, the solution scattering curve is unchanged by modification of the pH (range 6.0-8.0) or ionic strength (150-1500 mM NaCl) (Supplementary Figure 2.6) indicating the dynamic motion is an inherent property of the protein.

We can pinpoint the region responsible for the extraordinary range of motion in this protein, to an 11-residue peptide linking the trimerisation and catalytic domains (residues 39-KADEQQAQFRQ-49) (Figure 2.3). The 31 crystallographically characterised PmScsC protomers have the same thioredoxin-fold catalytic domains (Figure 2.4A and B) (RMSD 0.2-1.2 Å for 129 Cα). Similarly, all 31 PmScsC protomers share the same trimeric right-handed coiled coil stalk formed from the N-terminal residues (Figure 2.3A-C). At the secondary structure level, it is only the 11-residue linker peptide that changes significantly across the crystal structures (Figure 2.4C, Supplementary Movie 2.1).

The linker is a shape-shifting peptide that can adopt helical, strand or loop conformations. In the 24 protomers present in the compact PmScsC crystal structure, the linker forms a loop that positions the three catalytic domains close to the trimerisation stalk (Figure 2.3A and 4C). In the extended crystal structure the linker is helical, rotating and translating the catalytic domain away from the trimerisation domain relative to the compact structure (Figure 2.3C and 2.4C). In the transitional PmScsC crystal structure, each of the two trimers incorporates protomers in three conformations: compact, extended-like and intermediate (Figure 2.3B). The two compact protomer conformations of the transitional structure are similar to those in the compact crystal structure (RMSD 0.9-1.8 Å for 188 Cα). The two extended-like protomer conformations of the transitional structure are similar to the protomer in the extended crystal structure though they bend at different points in the linker helix, giving rise to somewhat different catalytic domain placements (Figure 2.4C). The two intermediate conformations in the transitional crystal structure have a linker that forms a short β-strand, and positions the catalytic domain directly above the trimerisation stalk (Figure 2.4C).

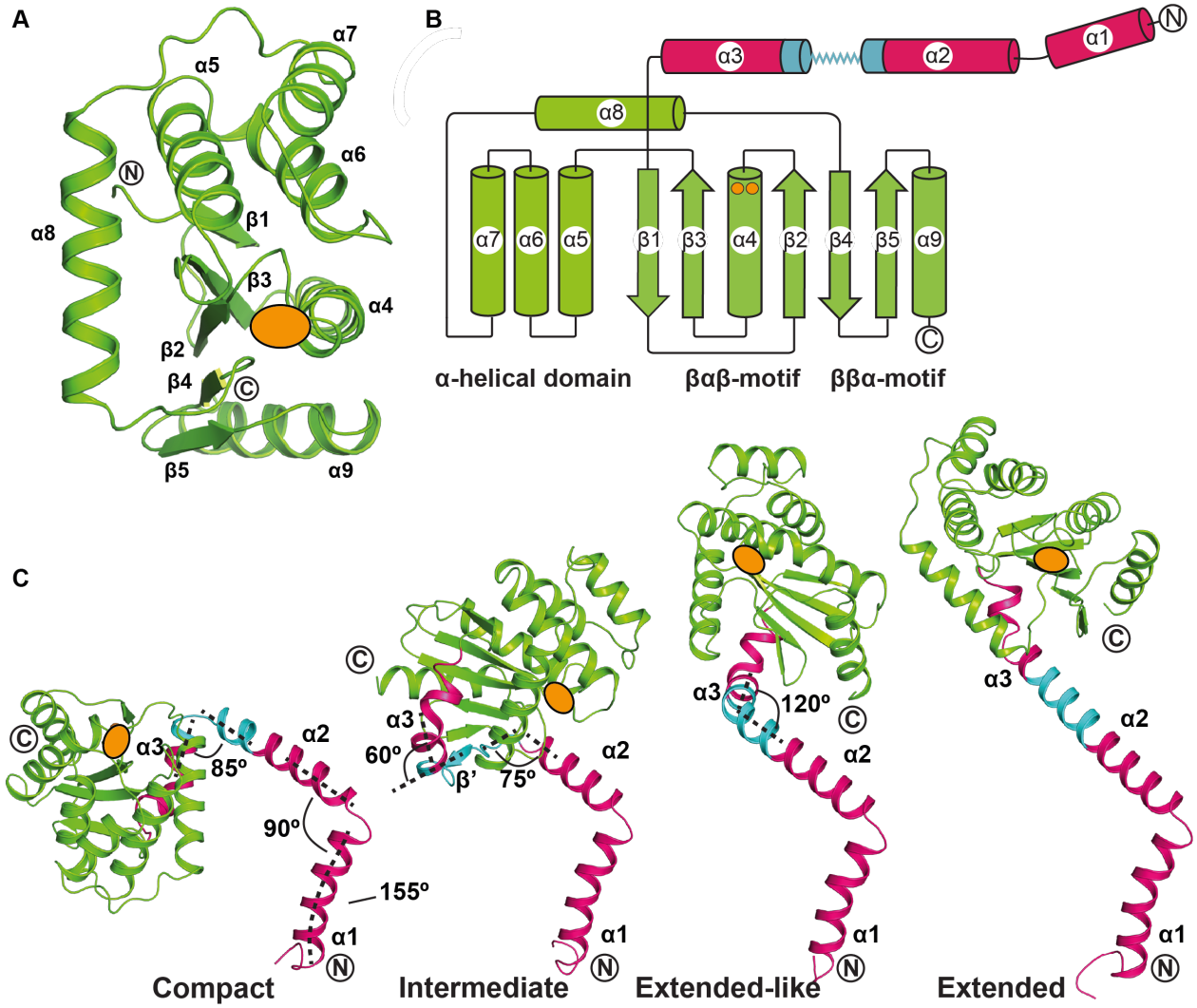


Figure 2.4: PmScsC protomer structure. A. *Cartoon representation of the PmScsC catalytic domain (equivalent to a DsbA fold, green). Secondary structure elements and N- and C- termini are labelled. Position of the active site CXXC motif is shown as an orange oval shape.* ***B****. PmScsC topology diagram showing the N-terminal helices α1/2/3 that precede the catalytic domain (magenta, flexible linker shown in cyan), and the DsbA domain (green) and CXXC active site (orange dots). Secondary structure elements and N- and C- termini are labelled.* ***C.*** *The PmScsC protomers are structurally diverse: compact, intermediate, extended-like and extended. Colour scheme is as for panel b. N-terminal helices α1/2/3 and angles between the helices are labelled. N- and C-termini are also labelled.*

As a consequence of the linker peptide flexibility, the three crystal structures describe extraordinarily different PmScsC quaternary shapes and dimensions. The 8 compact PmScsC trimers are mushroom-shaped with dimensions 80 Å × 80 Å × 50 Å (Figure 2.3A). The two intermediate PmScsC crystal structures are asymmetric flat triangles with dimensions 105 Å × 40 Å × 70 Å (Figure 2.3B). The extended PmScsC crystal structure is three-leaf-clover shaped with dimensions 75 Å × 80 Å × 80 Å (Figure 2.3C).

We hypothesised that the dynamic motion suggested by the crystal structures and supported by solution data is critical to the mechanism of PmScsC. Indeed, replacement of the shape-shifting peptide with a rigid helical peptide (PmScsC RHP, 39-AEAAAKEAAAKA-50) (135) reduced the flexibility of the protein as evidenced by small angle scattering data (Supplementary Figure 2.7), and abolished activity in the isomerase assay (Figure 2.1C). Furthermore, we showed that trimerisation of PmScsC is critical for activity because removal of the N-terminal residues (PmScsCΔN) also abolishes *in vitro* disulfide isomerase activity (Figure 2.1C) and *in vivo* complementation assays (swarming in the presence of copper) (Figure 2.1B).

Curiously, nature has also performed the ScsCΔN experiment. PmScsC shares 58% sequence identity with *Salmonella enterica* serovar Typhimurium StScsC (120) (Supplementary Figure 2.8) and both proteins contribute to copper resistance. Whereas we showed above that PmScsC is required for swarming motility under copper stress, StScsC is required for growth of *S.* Typhimurium in rich media in the presence of copper. By contrast, deletion of *scsC* in *P. mirabilis* had no effect on growth in rich media whether copper was present or absent (Supplementary Figure 2.2C). These findings imply that the two proteins have different target substrates and/or different molecular functions. Indeed, StScsC and PmScsC are very different proteins. StScsC lacks the N-terminal residues of PmScsC, it is monomeric rather than trimeric, and it has no isomerase activity. Intriguingly, the catalytic domains of the two proteins (as assessed by crystal structures) are very similar (PDB: 4GXZ) (RMSD 1.0-1.6 Å for 129 Cα). On the basis of this comparison, we predicted that removal of the N-terminal trimerisation residues from PmScsC would not only convert the trimer into a monomer and abolish isomerase activity, it might also convert the enzyme into a dithiol oxidase. Indeed, we confirmed that this was the case (Supplementary Figure 2.1D).

These two proteins, trimeric disulfide isomerase PmScsC and monomeric dithiol oxidase StScsC, are encoded in similar loci and are both associated with copper resistance. Yet, they have very different architectures, and different molecular and cellular functions. To examine if the region encoding the N-terminal extension of PmScsC is broadly conserved in *P. mirabilis*, the *scsC* gene was PCR amplified from 25 randomly selected clinical isolates and sequenced. We found that the *scsC* sequence was conserved in all 25 isolates – including the region encoding the trimerisation domain and shape-shifting motif - except for a single synonymous mutation in the signal sequence. This high degree of conservation suggests a highly conserved functional role. *Caulobacter crescentus* ScsC also has an extended N-terminal region and is predicted to be dimeric (118). Together with our work, this suggests that the N-terminal region of this protein is necessary for oligomerisation and that there are at least two very different molecular and functional classes of ScsC.

How does the trimeric and highly dynamic molecular architecture of PmScsC compare with other structurally characterised disulfide isomerases? Until this point, highly efficient disulfide isomerases were thought to function through the presence of two TRX domains (each with a CXXC active site) that can simultaneously interact with two substrate cysteines (51). The two active sites of V-shaped DsbC and U-shaped PDI are embedded within a dynamic framework that enable shuffling of substrate disulfides (47,136-138) (Figure 2.5A and B). Trimeric PmScsC also has the potential to interact with multiple substrate cysteines simultaneously. However, the evidence presented here shows a much more expansive range of motion than observed previously for DsbC or PDI, as depicted in morphing movies (Supplementary Movies 2.2-2.4) and static structural comparison (Figure 2.5C). For example, the distance between active sites of DsbC (PDB: 1JZD-1EEJ) (39,47) or PDI (PDB: 4EKZ-4EL1) (137) in reported crystal structures ranges between 31-39 Å (26% increase) and 28-40 Å (43%), respectively. By comparison, this distance ranges from 16-52 Å (225%) in PmScsC.

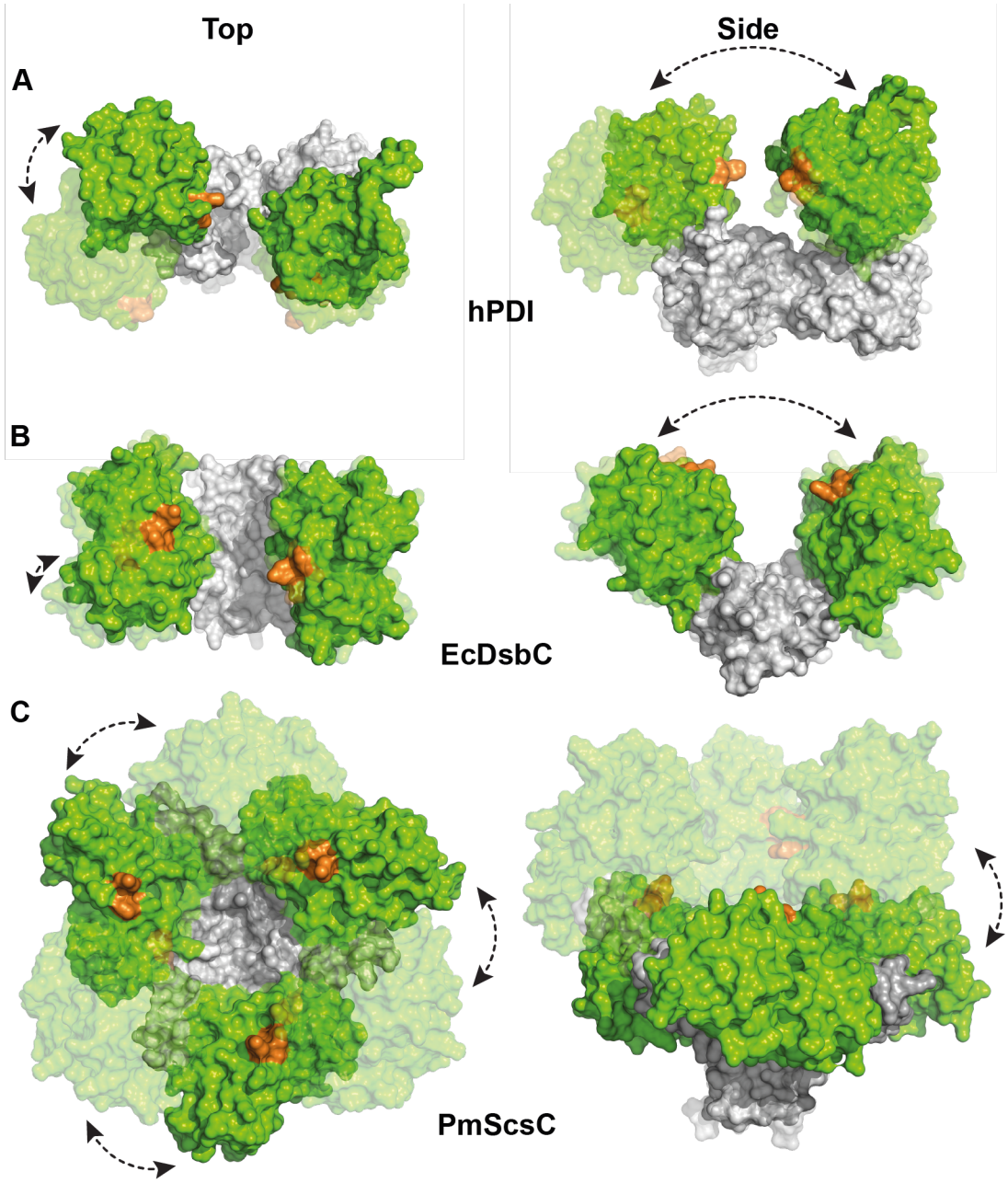


Figure 2.5: Disulfide isomerase range of motion. *Top and side views of* ***A.*** *U-shaped human PDI with two catalytic domains (PDB: 4EKZ and 4EL1),* ***B.*** *V-shaped homodimeric EcDsbC with two catalytic domains (PDB: 1JZD and 1EEJ) and* ***C.*** *homotrimeric PmScsC (this work, PDB: 4XVW and 5ID4) showing its expansive twisting and elongation motions. Catalytic domains are shown in green, active sites (CXXC) in orange, and oligomerisation or non-catalytic domains in white. See also Supplementary Movie 1.*

2.5 Discussion

Why is it that other characterised disulfide isomerases do not have the highly flexible, trimeric structure of PmScsC? How does the monomeric *S.* Typhimurium ScsC – not a disulfide isomerase – confer copper resistance? How do the very diverse ScsC proteins (monomeric, dimeric, or trimeric) form redox relay systems with a predicted partner protein ScsB (118) that appears to be highly conserved in bacterial Scs gene clusters. The likely explanation for the diversity in architecture and flexibility of ScsC proteins may relate to differences in their target specificity. The structural diversity among ScsC homologues may simply reflect the fact that each ScsC homologue has evolved to optimise disulfide reduction and/or isomerisation of a distinct protein substrate or set of substrates that are sensitive to copper-induced damage. Identifying and characterising these specific substrates, and understanding why these are not refolded or reduced by other periplasmic TRX-fold proteins are key questions that will inform our understanding of bacterial copper resistance mechanisms.

In summary, PmScsC is a unique and highly dynamic disulfide isomerase. The N-terminal residues bring three catalytic domains together into a trimer, and a unique flexible linker enables extraordinary twisting and extending motions that impact on the catalytic domain placement. Crystal structure snapshots supported by X-ray scattering data from solution, indicate an almost doubling in length of the protein in one dimension, in concert with considerable catalytic domain rotation. The combination, extent and diversity of these motions would enable mis-folded substrates bound to PmScsC to explore a broad folding landscape, consistent with the redox foldase activity of this key *P. mirabilis* copper resistance protein. These findings show in much greater detail than ever before how redox proteins with multiple active sites are able to shuffle incorrectly folded substrates. It is unclear whether PmScsC directly interacts with copper or how it interacts with other proteins required for virulence under copper stress. Nevertheless, the structural data may provide a suitable basis for drug discovery targeting a central defense mechanism of an important human pathogen (31,32). Importantly, the shape-shifting motif may also represent a modular component for dynamic motion that could be used in “plug and play” protein engineering applications.

2.6 Methods

2.6.1 Bacterial strains and growth conditions

BLAST searches of DsbA homologues were performed on the genome of *Proteus mirabilis* HI4320 (126). Clinical isolates of *P. mirabilis* were cultured from the urine or blood of patients with urinary tract infection or bacteremia. The isolates were sourced from the Princess Alexandra Hospital or Sullivan Nicolaides Pathology (Brisbane, Australia). *E. coli* Top10 (Invitrogen) was used for plasmid manipulations. *E. coli* and *P. mirabilis* were cultured in Luria-Bertani (LB) broth (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl) on LB agar (LB medium containing 15 g L^{-1} agar). To prevent swarming of *P. mirabilis*, LB medium containing 30 g L^{-1} agar and without NaCl was used. Media were supplemented with chloramphenicol (30 µg mL^{-1}), ampicillin (100 µg mL^{-1}) or kanamycin (25 µg mL^{-1}) as required.

2.6.2 Construction of *Proteus mirabilis scsC* mutants

Mutation of the *scsC* gene in *P. mirabilis* PM38 and PM54 was performed using the TargeTron® gene knockout system (Sigma-Aldrich) as per the manufacturer's instructions. Briefly, optimal intron insertion sites and primer sequences (5827, 5'-aaaaaagcttataattatccttaagtatcgaagatgtgcgcccagatagggtg, 5828, 5'-cagattgtacaaatgtggtgataacagataagtcgaagatgctaacttacctttctttgt, 5829, 5'-tgaacgcaagtttctaatttcgattatacttcgatagaggaaagtgtct) were predicted using the Sigma TargeTron online algorithm, followed by a retargeting PCR and cloning of the amplicon into the pACD4-K-C shuttle vector, resulting in the generation of plasmid pACD4-K-C*scsC.* Plasmid pAC4K-C*scsC* and pAR1219 (helper plasmid) were transformed into PM38 and PM54; induction and retargeting of the intron containing the kanamycin-resistance gene cassette into *scsC* was performed according to the manufacturer's instructions. Transformants with the correct intron insertion were selected by growth in the presence of kanamycin. Mutants with the correct insertion were confirmed by PCR and nucleotide sequencing using primers 5914 (5'-atgaaaaaaatttggctggcgttagc) and 5915 (5'-ttatttttttcatcagtaattgattgacgacatcagaataag), and referred to as PM38*scsC* and PM54*scsC*, respectively. The reprogrammed intron was inserted between nucleotide 528 and 529 in the sense strand of *scsC*. Plasmids pACD4-K-C*scsC* and pAR1219 were cured by passaging on nonselective medium followed by screening for loss of chloramphenicol and ampicillin resistance. Plasmid pScsC was constructed by amplifying the *scsC* gene from PM54 using primers PMscsC-pSU2718F (5'- gcgtctagattaactattctttcagaggctaaaggagcc) and PMscsCpSU2718R (5'-cgaagcttcgttatttttttcactttcgccagttgttctttcac), digestion with XbaI-

HindIII, and ligation into XbaI-HindIII digested pSU2718. Mutation of the CXXC active site (C82A; plasmid pScsC AXXC) and deletion of residues 1-42 (plasmid pScsCΔN) was performed by PCR; all constructs were confirmed by sequencing. These mutants retained the periplasmic signal sequence. Complementation was performed by introducing plasmids pScsC, pScsC AXXC, pScsCΔN, or pSU2718 (vector control) into PM54*scsC*.

2.6.3 Swarming motility in the presence of $CuSO_4$

Overnight cultures of *P. mirabilis* were diluted in fresh LB broth to an OD_{600} of 0.1. Swarming on LB agar was examined in the absence or presence of $CuSO_4$ (final concentration of 1.5 mM). For assays performed in the presence of $CuSO_4$, a 10 µL sample of the diluted *P. mirabilis* PM54 and PM54*scsC* strains, respectively, was inoculated onto the center of an agar plate and allowed to soak into the agar (resulting in equal diameter for each sample). The plates were incubated at 37 °C. The diameter of the swarming zone was measured after 14 h incubation and every one or two hours thereafter for a total of 10 h. Swarming of PM54*scsC* containing pSU2718 (vector control) or complementation plasmids (pScsC, pScsC AXXC and pScsCΔN) was performed as described above but with the addition of chloramphenicol (30 µg mL^{-1}) and 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for ScsC induction. Swarming motility at 37 °C on LB agar was measured after 6 h and 8 h of incubation. Three replicate experiments were performed for all assays. Swarming motility of PM38 wild-type, mutant and complemented strains in the presence and absence of copper was performed in the same manner.

2.6.4 Western blotting

ScsC expression in *P. mirabilis* wild-type, mutant and complemented strains was examined by western blot analysis. Crude cell lysates were prepared from standardised overnight cultures. Samples were resolved by SDS-PAGE and transferred onto a polyvinylidene difluoride membrane. Blots were probed using a PmScsC specific polyclonal antiserum raised in rabbits (WEHI Antibody Facility) at 1:1000 dilution in phosphate buffer saline with 0.05% *v/v* Tween-20.

2.6.5 Protein production

Codon-optimised PmScsC (see Supplementary Note 1 for the sequence) lacking the first 21 amino acids (which correspond to the predicted secretion signal) or PmScsCΔN (lacking the first 63 residues, starting at $A^1Q^2F^3$…) was inserted into pMCSG7 (Midwest Center for Structural Genomics) by ligation-independent cloning. PmScsC and PmScsCΔN

were expressed in BL21(DE3)pLysS cells (Novagen) at 30ºC for 16-20 h using ZYP-5052 autoinduction media (139). Selenomethionine-labelled PmScsC (mature PmScsC contains 4 native methionine sites) was expressed in BL21(DE3)pLysS cells grown in minimal media (M63) supplemented with 0.05 mg mL^{-1} D/L-selenomethionine. Recombinant protein expression was induced by 0.1 mM IPTG overnight at 30ºC.

After expression, all cells were harvested by centrifugation then resuspended in lysis buffer (25 mM Tris pH 7.4, 150 mM NaCl) with added protease inhibitor (1:1000 dilution into lysate, BioPioneer, Inc., USA) and DNase I (3.3 µg mL^{-1} final concentration, Sigma-Aldrich). The bacterial cells were processed through a Constant Systems (LTD, UK) TS-Series cell disruptor twice, at pressures of 22 and 24 kpsi. The lysate was clarified by centrifugation (40 000 × *g*, 30 min, 4 °C, rotor JA-25.5, Beckman Coulter, Brea, CA) and the soluble fraction was run twice over 2×3 mL of equilibrated Talon® (Clontech) resin to bind the His-tagged protein. The resin was washed with 6 column volumes of lysis buffer before the protein was eluted with 25 mM Tris pH 7.4, 150mM NaCl, 250 mM imidazole. To remove the His_6-tag, Tobacco Etch Virus (TEV) protease cleavage was performed at room temperature for 2 hours with 1 mg of TEV protease per 50 mg of His-tagged protein. The protein mixture was then desalted using a Sephadex® G-25 fine 16/60 column connected to an ÄKTA™ FPLC system (GE Healthcare) and reverse IMAC (2×2 mL of Talon resin) was used to separate cleaved protein from the His-tagged contaminants. PmScsC and PmScsCΔN were reduced or oxidised with a 25-fold molar excess of DTT, or a 10-fold molar excess of copper(II)/1,10-phenanthroline, respectively. The final step of purification for the proteins was size-exclusion chromatography using a Superdex® S75 column with 10 mM HEPES pH 7.4, 150 mM NaCl. Proteins were concentrated using Amicon® Ultra centrifugal filter devices with a 10-kDa cutoff (Merck Millipore, USA). Yields for PmScsC derivatives were around 80 mg per litre of culture. SDS-PAGE (NuPAGE® 4-12% BisTris gel, Invitrogen, Australia) with Coomassie blue stain was used to assess the protein quality. Concentration of protein samples was determined using the A_{280} of the sample (read using a Thermo Scientific NanoDrop™ 2000c spectrophotometer) and calculated extinction coefficients from ProtParam (140). *E. coli* DsbA and *E. coli* DsbC, lacking the periplasmic leader signal were expressed and purified as described for PmScsC.

2.6.6 Design of PmScsC rigid helical peptide mutant

The 11-amino-acid flexible linker peptide in PmScsC (residues 39-KADEQQAQFRQ-49) was replaced with a 12-amino-acid rigid helical peptide (39-AEAAAKEAAAKA-50) (135) using overlap extension PCR (141) with the primers, 5'-gcagaagcagcggccaaagaggctgcggctaaagccgcactggctagcgaacatgatgcc and 5'-ggctttagccgcagcctctttggccgctgcttctgctttcgtctgcagagccatgattgc. The mutated gene was then reinserted into pMCSG7 by ligation-independent cloning. The construct was confirmed by sequencing and the protein expressed and purified as described for wild-type PmScsC.

2.6.7 PmScsC sequence conservation

All clinical isolates were de-identified, and individual informed consent was not required. Genomic DNA isolation was performed using the UltraClean® Microbial DNA Isolation Kit (MO-BIO Laboratories, #12224) according to the manufacturer's instructions. Yields varied between 100-500 ng μL^{-1} of DNA. PCR amplification of the *scsC* gene from *P. mirabilis* strains was performed using a sequence specific primer pair (5'-gtgccgtttaaccagatttatg) and (5'-cgtagataaatcagtaagttctg) in combination with Phusion® High-Fidelity DNA Polymerase (New England BioLabs Inc., MA, USA). DNA was initially denatured in a single step at 98ºC for 30 s followed by 30 cycles of denaturation (98ºC, 10 s), annealing (50ºC, 20 s) and elongation (72ºC, 30 s). In a final step fragment elongation was completed at 72ºC for 7 min. PCR samples were examined by electrophoresis and subsequent DNA sequencing.

2.6.8 Chemical cross-linking

150 µM of purified PmScsC (in 50 mM HEPES buffer pH 8.5) was reacted with 3 mM homobifunctional dithiobis(sulfosuccinimidylpropionate) (DTSSP) (Pierce) for different time intervals (30 sec – 30 min) at 23ºC. DTSSP contains one amine-reactive *N*-hydroxysulfosuccinimide (sulfo-NHS) ester at each end of its eight carbon spacer arm. The sulfo-NHS group reacts preferentially with amines (from lysine residues) between pH = 7-9, to form a stable amide bond. DTSSP contains a disulfide bond in the spacer arm, which can be readily cleaved with dithiothreitol (DTT). Aliquots were removed at various time points and mixed with $(NH_4)HCO_3$ (50 mM final), to stop the reaction, through hydrolysis of remaining functional ester groups in DTSSP. The samples were then run on a non-reducing SDS-PAGE and subjected to Coomassie staining. As control, 50 mM DTT was added to selected samples, which led to cleavage of the internal disulfide bond in DTSSP, and consequently disruption of cross-links between PmScsC protomers.

2.6.9 MALLS

A combined approach, using analytical size exclusion chromatography (SEC) and multiangle laser light scattering (MALLS) was utilised to determine and compare the stoichiometry of PmScsC with PmScsCΔN in solution. The setting consisted of an LC20 high-performance liquid chromatography (HPLC) system (Shimadzu, Rydalmere, Australia) and a DAWN® HELEOS® II laser light detector connected to an Optilab® T-rEX™ refractive index detector (Wyatt Technology, Dernbach, Germany). Either a Superdex 200 or Superdex 75 10/300 GL analytical column (GE Healthcare, USA) was connected to the LC20 HPLC system and equilibrated with 25 mM Tris and 150 mM NaCl, pH 7.5 overnight. Purified proteins were injected (500 µL of 3 mg mL^{-1} at a flow rate of 1.0 mL min^{-1}) into the SEC-MALLS system for analysis. To calibrate the detector, 500 µL of 5 mg mL^{-1} bovine serum albumin (BSA) (Sigma-Aldrich, Australia) was used in 25 mM Tris and 150 mM NaCl, pH 7.5 at a flow rate of 1.0 mL min^{-1}. Wyatt Astra V software was used for data collection and analysis.

2.6.10 SAXS

Small angle X-ray scattering data were collected on the SAXS-WAXS beamline at the Australian Synchrotron (142). Immediately prior to loading, all samples were centrifuged at 10,000 *g* to remove large particles from the solution, and radiation damage was minimised by flowing samples (~100 µL) past the beam in 1.5 mm quartz capillaries (Hampton Research). Data reduction was carried out using the Australian Synchrotron Scatterbrain software (143) correcting for sample transmission and solvent scattering.

Scattering data were collected on reduced PmScsC wildtype and RHP mutant proteins in 25 mM HEPES pH 7.5, 150 mM NaCl, 1mM DTT. Data were also collected on oxidised wildtype PmScsC, but as no significant differences in the scattering were observed, only data from the reduced forms are presented. To confirm that the pH and ionic strength of the crystallisation conditions do not induce conformational changes in PmScsC, scattering data were collected from wildtype PmScsC prepared at 0.50 mg mL^{-1} in a gradient of 25mM HEPES pH (6.0, 6.5, 7.0, 7.5, and 8.0) and NaCl (150, 300, 600, 900, and 1500mM). No significant systematic change in the structural parameters was observed at any point of the gradient, hence, it was concluded that neither ionic strength nor pH influence the conformation of the complex in the ranges measured.

Data quality was assessed by inspection of the linearity of the Guinier region of the data (qR_g < 1.3), estimated molecular mass of the protein complex, and concentration dependence of the scattering. The estimated molecular mass was determined as outlined in (144), where the contrast and partial specific volume were estimated from the protein sequence (145). The pair-distance distribution function ($p(r)$) was generated from the experimental data using *GNOM* (146) from which $I(0)$, R_g and D_{max} were determined. Rigid body modelling of the scattering data from PmScsC wildtype and mutant proteins was performed using *CORAL* (147). C_3 symmetry was assumed, and the starting model was oriented such the 3-fold axis was parallel with the z-axis, and passed through the centre of the oligomerisation domain. Two rigid bodies were then defined for each monomer: residues 3-44 (oligomerisation domain); 47-224 (catalytic domain). The position of the oligomerisation domain was fixed, and the position and orientation of the catalytic domain was then optimised against the measured scattering data. The program was run 16 times for each protein, but the models with the lowest penalty function still showed small systematic deviations from the experimental data (Wildtype: χ^2=3.4; CorMap test (148), 182 points, C = 51, P = 0.000; RHP mutant: χ^2=5.6; CorMap test, 171 points, C = 28, P = 0.000). Given the crystal structures show significant structural diversity, it is likely that these systematic deviations arise from the fact that there is an ensemble of structures present in solution. Hence, ensemble optimisation was also performed with the program EOM(149). From an initial pool of 1000 structures (the oligomerisation domain possessed C_3 symmetry in all structures, but the entire trimer was permitted to adopt either C_1 or C_3 symmetry), a final ensemble of 6 structures was obtained that yielded an excellent fit to the data (Wildtype: χ^2=1.0; CorMap test, 182 points, C = 12, P = 0.021; Mutant: χ^2=1.8; CorMap test, 171 points, C = 11, P = 0.076). The ensemble for the wildtype protein is composed of 6 structures and spans a diverse range of conformations. Thus, the SAXS data supports the notion that the PmScsC trimer is highly dynamic in solution. The ensemble for the mutant protein is composed of 4 structures, all in extended conformations. The initial pool of random structures was the same for both WT and RHP optimisations, and the number of structures in the final ensemble was not artificially limited. The reduction in size of the ensemble (from 6 for WT to 4 for RHP) is consistent with the RHP mutation rigidifying the flexible helix, and reducing the conformational space sampled by the protein. Details of the data collection and structural parameters are summarised in Supplementary Table 2.1.

2.6.11 Crystallisation and structure determination

Crystallisation screenings were performed at the UQ ROCX facility at the University of Queensland (uqrocx.imb.uq.edu.au) using commercial screens and the hanging-drop vapour diffusion method.

Compact structure

Selenomethionine-labeled PmScsC crystals were grown at 20 ºC from a drop comprising 200 nL of 30 mg mL^{-1} purified protein in 10 mM HEPES pH 7.4 and 200 nL of well solution, 2.85 M sodium malonate pH 5.8 containing 0.1 M cobalt(II)chloride hexahydrate. Crystals grew over a period of 2-3 days. A data set at the Se-edge (wavelength = 0.9792 Å) was measured under cryogenic conditions (100 K) at the Australian Synchrotron MX2 beamline using the BluIce interface (150). Data were processed in XDS (151) and scaled to space group $P2_12_12$ in AIMLESS (152). A manually guided molecular replacement approach was performed using StScsC as the template (PDB ID: 4GXZ) in MOLREP (153). Manual selection of solutions based on observed packing, allowed placing of six molecules in the asymmetric unit. This solution was improved in the subsequent MR runs in MOLREP (153), allowing the addition of eight molecules in the second trial and a total of nine molecules in the third run. A SAD-MR approach in PHENIX (154), by combining SAD phasing and the partial MR solution (9 molecules), and subsequent density modifications resulted in an electron density map with a skew value of 0.82 (1826 residues (190 side chain built) in 250 fragments). Successive AutoBuild runs in PHENIX (154) built 1557 residues (806 side chain) and dropped the R and Rfree values to 38% and 41.5%, respectively. Extensive manual building was required to complete one copy of PmScsC. Molecular replacement using this model as template in PHASER (155) allowed placing of 12 molecules in the asymmetric unit. After several rounds of manual building in Coot (156) and PHENIX (154) refinements R/Rfree values dropped to 29.1 and 32.3%, but further refinement cycles were unsuccessful. Careful examination of Ctruncate (157) indicated the presence of nearly perfect pseudo-merohedral twinning (-h,-k,l) and led to a reconsideration of space group. Reprocessing the data set to a resolution of 2.6 Å in space group $P2_1$ yielded lower R-merge (0.072 vs 0.097 for data processed in $P2_12_12$). A new molecular replacement run using the reprocessed data in *P*21 space group and the PmScsC template with a resolution cut off set to 3.2 Å finally identified 24 copies in PHASER (155). Rigid body refinement in REFMAC (158) showed that refinement considering the twin operator (-h,-k,l) produced better R-factors (37% vs 42%, no twinning). After manual correction of two chains that were placed inaccurately during

molecular replacement by PHASER (155), and several rounds of refinement in PHENIX (154) and REFMAC (158) the final R/Rfree values are 24.8/28.2% at 2.6 Å resolution. The Ramachandran favoured/outlier statistics for the structure are 95%/0.9%. A stereo image of representative electron density for this structure is show in Supplementary Figure 2.9A.

Transitional structure

Crystals were grown at 20 °C from drops containing 1 µL of 48 mg mL^{-1} oxidised PmScsC in 10 mM HEPES pH 7.4 and 1 µL of 2.85 M sodium malonate pH 5.8 containing 0.1 M copper (II) chloride. Crystals formed after 2-3 days and were cryoprotected with 3.4 M sodium malonate pH 5.8 and flash frozen in liquid nitrogen. Data were collected using the BluIce software (150) on the MX2 beamline of the Australian Synchrotron at a wavelength of 0.9792 Å under cryogenic conditions (100 K). The data were processed with XDS (151) and scaled in AIMLESS (152) using the autoPROC framework (159). The space group was determined to be *I*4 and molecular replacement using PHASER (155), as implemented in the CCP4 suite (160), was performed with a single catalytic domain from the compact structure (residues 46-224). Five molecules were initially found in the asymmetric unit and after assessing the crystal packing one of these was deleted and a subsequent round of PHASER (155) was run with the altered solution as an additional input model. This resulted in the placement of 6 molecules in the asymmetric unit, which is consistent with the presence of two trimers. Manual building of the trimerisation domain was performed in Coot and rounds of autobuild using Buccaneer (161) in CCP4 and refinement using REFMAC (158) and BUSTER (162,163) were performed to improve the structure. Several rounds of manual adjustment in Coot (156) and refinement in PHENIX (154) were performed to yield final R/Rfree values of 17.1% and 22.2%. Validation in MolProbity (164) was used throughout the refinement process, and the final Ramachandran favoured/outlier statistics for the structure are 98%/0.5%. A stereo image of representative electron density for this structure is show in Supplementary Figure 2.9B.

Extended structure

Crystals were grown at 20 °C from drops containing 1 µL of 20 mg mL^{-1} oxidised PmScsC in 10 mM HEPES pH 7.4, 150 mM NaCl and 2.5 mM copper (II) chloride (added to protein solution immediately before crystallisation) and 1 µL of 32% Jeffamine M-600 pH 7 in 0.1 M HEPES pH 8. Crystal formation occurred within minutes of set up and crystals continued to grow until day two. Crystals were cryoprotected with 32% Jeffamine M-600 pH 7, 0.1 M HEPES pH 8 and 20% ethylene glycol and frozen in liquid nitrogen. Data were collected

using the BluIce software (150) on the MX2 beamline of the Australian Synchrotron at a wavelength of 0.9537 Å under cryogenic conditions (100 K). Using the autoPROC framework (159) the data were processed with XDS (151) and scaled in AIMLESS (152). The Wilson B factor for the data was relatively high (~double that of the data for the other two crystal structures). The space group was determined to be *H*32 and molecular replacement using PHASER (155), as implemented in the CCP4 suite (160), was performed with a single catalytic domain from the compact structure (residues 46-224). Only one molecule was found in the asymmetric unit (the trimer is generated by crystal symmetry). Manual building of the trimerisation domain was performed in Coot with rounds of autobuild (Buccaneer (161)) and refinement (BUSTER (163)) performed to improve the structure. Several rounds of manual adjustment in Coot (156) and refinement in PHENIX (154), including the use of TLS, were performed to yield final R/Rfree values of 25.1% and 26.3%. Validation in MolProbity (164) was used throughout the refinement process and the final Ramachandran favoured/outlier statistics for the structure are 98%/0.5%. A stereo image of representative electron density for this structure is show in Supplementary Figure 2.9C.

Residues in all three crystal structures were numbered based on their position after the TEV cleavage site in the construct. MUSTANG-MR (165,166) was used for structure alignment and RMSD calculations and PyMOL was used to create images of the structures and perform other measurements. The distance range between active sites in DsbC, PDI and PmScsC structures was measured using the Cβ positions of the more N-terminal of the two catalytic cysteines. Supplementary Movie 1 was created in Pymol by generating morphs between each of the PmScsC crystal structures, as well as morphs between each of the available structures for *E. coli* DsbC (PDB: 1JZD-1EEJ) and human PDI (PDB: 4EKZ-4EL1).

2.6.12 Measurement of PmScsC redox potential and pK_a values

The redox potential of PmScsC was determined by incubating 2 μM PmScsC with 1 mM GSSG in combination with varying concentrations of GSH (4 -15 mM) for ~12 h at 24 °C in 100 mM sodium phosphate buffer, pH 7.0, and 1 mm EDTA. Trichloroacetic acid, 10% *v/v*, was used to precipitate the protein samples and the pellets were washed with ice-cold acetone. The samples were then treated with 2 mM 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid, in 50 mM Tris-HCl, pH 8.0, and 1% SDS, which blocks free thiols and increases the weight of the reduced protein by ~1 kDa. The reduced and oxidised forms of

PmScsC were separated on a 12% non-reducing SDS-PAGE (NuPAGE) and stained with Coomassie Brilliant Blue. The fraction of reduced protein, *R*, was determined via densitometric analysis of the stained SDS-PAGE gel using ImageJ (167), and the equilibrium constant $\mathbf{K}_{eq}$ was calculated via $\mathbf{R} = ([GSH]^2/[GSSH])/(\mathbf{K}_{eq} + ([GSH]^2/[GSSH]))$ (168). The Nernst equation, $\boldsymbol{E}^0 = \boldsymbol{E}^0_{GSH/GSSG} - (RT/\mathbf{nF})\cdot \ln \mathbf{K}_{eq}$, where $\mathbf{E}^0_{GSH/GSSG}$ is the standard potential of −240 mV (169), was used to calculate the redox potential of PmScsC.

The pH-dependent absorbance of the catalytic thiolate anion of PmScsC was determined at 240 nm (8) using a CARY 50 UV/VIS spectrophotometer (Agilent Technologies, USA). The pH titration measurements of oxidised or reduced PmScsC (40 µM) were conducted at 22 ºC in 2 mL reaction buffer (10 mM Tris, 10 mM sodium citrate, 10 mM K_2HPO_4, 10 mM KH_2PO_4, 200 mM KCl, and 1 mM EDTA). Absorbance (λ = 240 and 280 nm) was measured between pH 6.5 and 1.5 in 0.25 increments. The pK_a value was calculated from the fitted curves of three replicates using the Henderson-Hasselbalch equation (pH = pK_a - log ([A240/A280]red/[A240/A280]oxid)). Redox potential and pK_a experiments were each repeated three times.

2.6.13 Relative stability of oxidised and reduced PmScsCΔN

Unfolding of PmScsCΔN was monitored using the change in the far-UV circular dichroism (CD) signal (83). The largest difference in molar ellipticity for oxidised and reduced enzymes was calculated from initial far-UV CD spectra (from 250 nm to 190 nm) recorded at 25ºC and 95ºC using a Jasco J-810 circular dichroism (CD) spectropolarimeter (Jasco, USA), respectively. The unfolding of oxidised and reduced protein ($PmScsC\Delta N_{ox}$ = 222.5 nm, $PmScsC\Delta N_{red}$ = 219.5 nm) was monitored at a constant heat rate of 1 ºC min^{-1} starting from 25 ºC and increasing to 95 ºC in a 1 mm quartz cuvette. Unfolding reactions were performed using 10 µM protein in 100 mM NaH_2PO_4/Na_2HPO_4, 1 mM EDTA at pH 7.0. Reduced enzyme samples contained 0.75 mM DTT. The CD data was converted to a fraction of folded protein via $\alpha_{folded}(T) = (\theta(T) - \theta_{unfolded})/(\theta_{folded} - \theta_{unfolded})$ and $\alpha_{folded}(T)$ was then fitted using a Boltzmann sigmoid function $1 - 1/(1 + e^{N(Tm-T)})$, where *N* is a scaling factor that describes the nature of the transition and T_m is the protein melting temperature. The thermal unfolding experiment was repeated three times for both the oxidised and reduced protein.

2.6.14 Protein disulfide isomerase assay

The bovine pancreatic ribonuclease RNase A contains 4 disulfide bonds that are necessary for its enzymatic function. Starting with “scrambled” RNase A (in which the cysteines have formed non-native disulfide bonds), one can determine the ability of other enzymes to restore its function by spectroscopically monitoring the hydrolysis of 3’,-5’-cyclic monophosphate (cCMP4), a synthetic substrate of the ribonuclease. Thus, *in vitro* disulfide isomerase activity of PmScsC was monitored by using scrambled RNase A as substrate (170). Inactive scrambled RNase A was produced by treating 70 mg of native RNase A (Sigma-Aldrich) with 6 M guanidinium chloride and 150 mM DTT in 50mM Tris pH 8, incubated at room temperature for 16 h. After adjusting the solution to pH 8, the unfolded protein was then desalted and buffer exchanged into 100 mM acetic acid/NaOH pH 4. The presence of reduced cysteines was confirmed spectrophotometrically by mixing an unfolded RNase A sample with 8-fold higher molar concentration of 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB or Ellman’s reagent), which reacts with thiols to absorb at 412nm. The unfolded RNase A was then air-oxidised over 5 days in 6M guanidinium chloride, 50 mM Tris pH 8.5 at room temperature in the dark. After concentration of the protein and readjustment of the pH to 8.5, the scrambled RNase A was buffer exchanged into 100 mM acetic acid/NaOH pH 4 and concentrated to 11.5 mg/ml for use in assays. The absence of reduced cysteines was confirmed as described above. The total reaction volume for the scrambled RNase A assay was 750 µL containing 100 mM sodium phosphate buffer pH 7.0, 1 mM EDTA, 8.2 µM DTT, and 10 µM PmScsC, PmScsCΔN, PmScsC RHP, EcDsbC or EcDsbA and 40 µM of scrambled RNase A. At various time intervals, 50 µL sample was removed and mixed with 150 µL of 3 mM cytidine 3’,-5’-cyclic monophosphate (cCMP) to measure RNase A activity. Hydrolysis of the cyclic phosphate ester bond in cCMP was monitored at 296 nm using a Synergy H1 multimode plate reader (BioTek, USA). Samples containing native RNase A and scrambled RNase A in the absence of any other enzyme served as positive and negative controls, respectively. The disulfide isomerase assays were repeated twice for each protein.

2.6.15 PmScsC dithiol oxidation activity

The potential of PmScsC/PmScsCΔN to catalyse disulfide bond formation *in vitro* was assessed using a synthetic model peptide (CQQGFDGTQNSCK) as described before (171). Dithiol oxidation was determined fluorometrically, using a Synergy^{TM} H1 multimode plate reader (BioTek® Instruments, USA) with the excitation wavelength set to 340 nm and emission to 615 nm. A 150 µs delay before reading and 100 µs reading time were used for

time-resolved fluorescence. The reaction buffer contained 50 mM MES, 50 mM NaCl and 2 mM EDTA at pH 5.5. The reaction volume was 50 µL in each well of a white 384-well plate (Perkin Elmer OptiPlate^{TM}-384, Part #: 6007290), including increasing concentrations (80-320 nM) of EcDsbA, EcDsbC, PmScsC, or PmScsCΔN and 2 mM glutathione (GSSG) as oxidant. Lastly 8 µM peptide substrate was added to initiate the reaction. EcDsbA served as positive control. As negative control, reactions in the absence of proteins were used. The initial linear portion of the raw data was used to calculate the rate of oxidation as fluorescence increase per minute. The dithiol oxidation activity experiments were repeated three times for each protein.

2.7 Data availability

The UniProt accession code for the *P. mirabilis* ScsC protein described in this paper is C2LPE2 and the accession codes for other proteins described in this study are H9L4C1 (StScsC), P0AEG4 (EcDsbA), and P0AEG6 (EcDsbC). Coordinates and structure factors for the three ScsC crystal structures were deposited in the protein data bank (PDB) with accession codes 4XVW (compact), 5IDR (transitional) and 5ID4 (extended). In addition to the PDB accession codes listed above, PDB accession code 4GXZ (StScsC) was used as a molecular replacement template and 1EEJ and 1JZD (EcDsbC), and 4EKZ and 4EL1 (hPDI) were used to generate Figure 2.5. Scattering data and models have been deposited in SASBDB (172) with accession codes SASDB94 (wild-type PmScsC) and SASDBW6 (RHP mutant). All other data are available from the corresponding authors upon reasonable request.

2.8 Acknowledgements

We acknowledge use of the UQ ROCX Diffraction Facility and of the MX1, MX2 and SAXS/WAXS beamlines at the Australian Synchrotron. We thank the support teams at both facilities for advice and expert assistance and we are grateful to Gordon King for chemical cross-linking support. This work was supported by the Australian Research Council through a Laureate Fellowship (FL0992138, JLM), Future Fellowships (FT100100662 MAS; FT130100580, BH) and DECRA Fellowship (DE130101169, MT); and the National Health and Medical Research Council (Senior Research Fellowships to JLM, GNT455829 and to MAS, GNT1106930; project grant to MAS, GNT1106590).

2.9 Additional information

Competing interests: The authors declare no competing financial interests

2.10 Supplementary information

2.10.1 Supplementary Tables

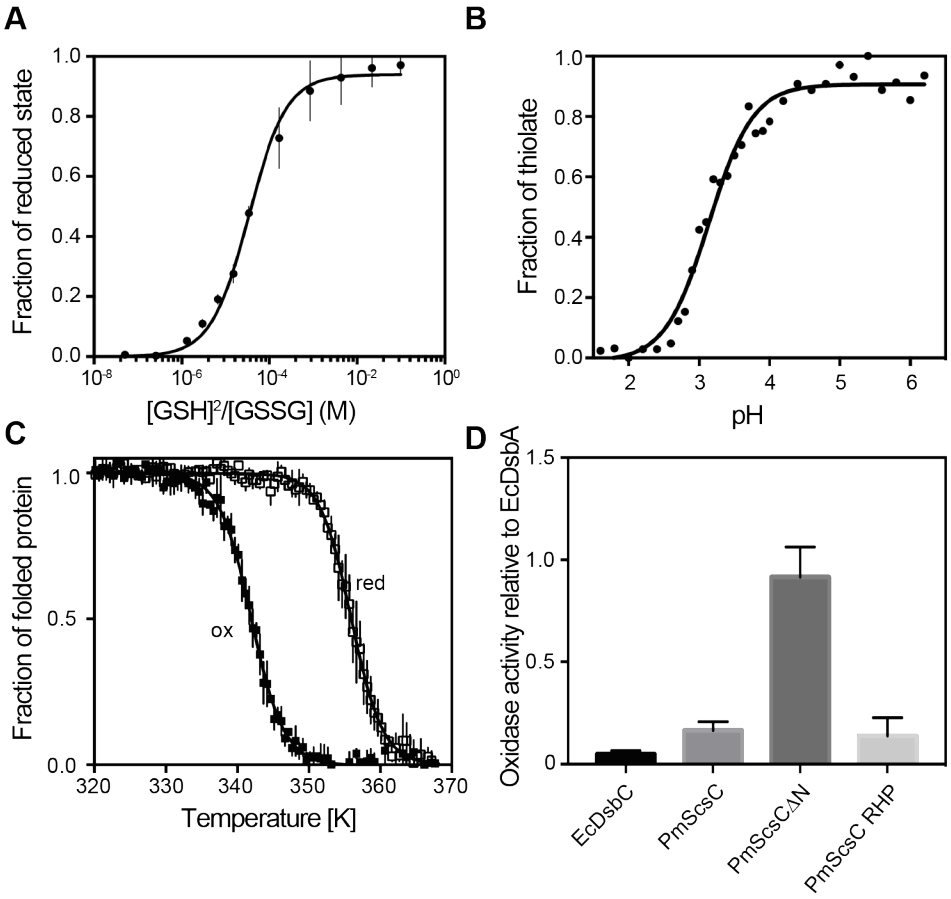
Supplementary Table 2.1: SAXS data collection and analysis details.

Data collection parameters	**PmScsC (*wild type*)**	**PmScsC (*rigid mutant*)**
Instrument	SAXS-WAXS (Australian Synchrotron)	
Beam geometry	Point	
Wavelength (Å)	1.033	
Camera length (m)	1.576	1.485
q-range (Å^{-1})	0.01-0.55	0.01-0.55
Exposure time (s)	28 (14 × 2 s exposures)	23 (23 × 1 s exposures)
Protein concentration (mg/mL)	2.15	1.30
Temperature (°C)	10	10
Normalisation standard	Water	Water
Structural parameters		
$I(0)$ (cm^{-1}) [from $p(r)$]	0.1006 ± 0.0002	0.06648 ± 0.00007
R_g (Å) [from $p(r)$]	36.0 ± 0.1	44.1 ± 0.1
$I(0)$ (cm^{-1}) [from Guinier]	0.1010 ± 0.0005	0.0663 ± 0.0001
R_g (Å) [from Guinier]	36.5 ± 0.3	43.8 ± 0.2
D_{max} (Å)	105 ± 5	135 ± 7
Porod volume (Å^3)	97800 ± 9000	108000 ± 10000
R_g (Å) [crystal structures]	30.5 – 35.8	
D_{max} (Å) [crystal structures]	90 – 105	
Dry volume (Å^3) [from sequence]	91000	
Molecular mass determination		
Partial specific volume (cm^{3}g^{-1})	0.73	
Contrast, $\Delta\rho$ (10^{10} cm^{-2})	2.75	
Molecular mass M_r [from $I(0)$]*	69900 ± 3500	76400 ± 3500
Molecular mass M_r [from sequence]	74100	73300
Software employed		
Primary data reduction	*Scatterbrain*	
Data processing	*PRIMUS and GNOM*	
Rigid body modelling	*Coral and EOM*	

*For the wild type protein, data were also collected at protein concentrations of 4.50 and 8.85 mg/mL. The corresponding $I(0)$ values for these concentrations are 0.2120 and 0.4060 cm^{-1}. These values yield molecular mass estimates of 70.3 and 68.6 kDa. Thus, there is no evidence of significant interparticle interactions in the measured concentration range. For the rigid mutant protein, data were also collected at protein concentrations of 0.60, 2.65 and 5.30 mg/mL. The corresponding $I(0)$ values for these concentrations are 0.0312, 0.1447 and 0.2940 cm^{-1}. These values yield molecular mass estimates of 77.9, 81.6 and 82.9 kDa. Thus, there is evidence of weak interparticle interactions in the

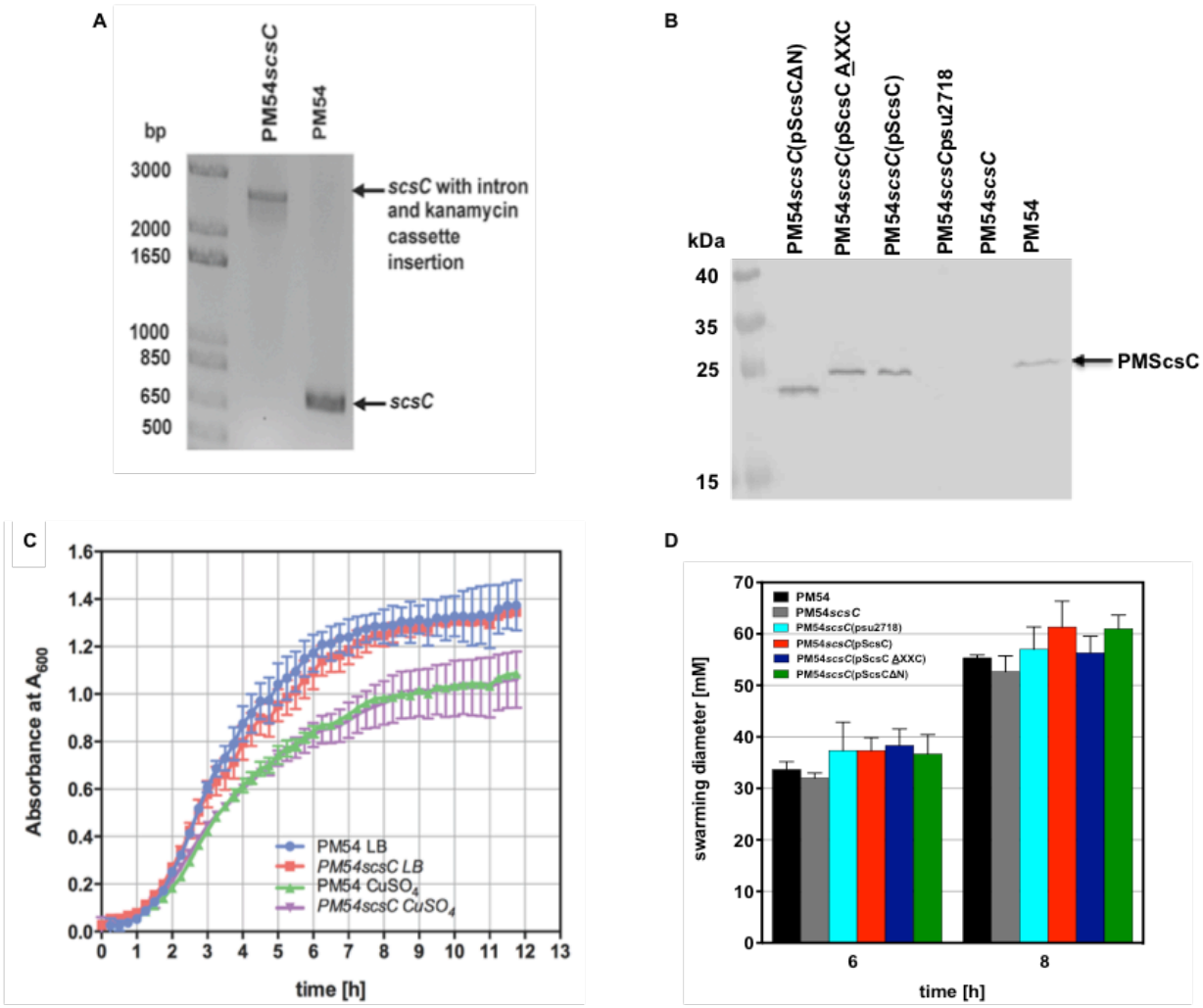
measured concentration range that is deemed negligible at the concentration used for data modelling and analysis.

2.10.2 Supplementary Figures



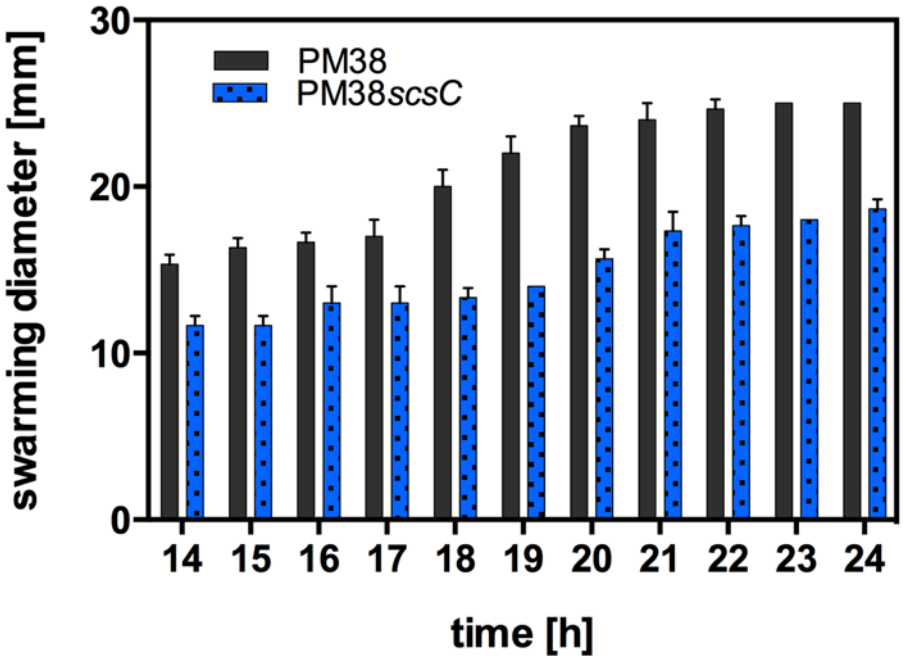
Supplementary Figure 2.1: Functional properties of PmScsC. A. *Redox potential. PmScsC was equilibrated in a range of glutathione (GSSG/GSH) redox buffers to measure the equilibrium constant K_{eq} (34 µM, which corresponds to a redox potential of -108 mV). Data are shown as the mean ± s.d., of three replicate experiments.* ***B.*** *Determination of the pK_a of the highly reactive cysteine in the CXXC active site motif. Presence of the catalytic thiolate anion is pH-dependent and can be spectroscopically monitored. The pK_a of PmScsC CXXC was determined to be 3.1. A representative example of the three replicate experiments is shown* ***C.*** *Thermal unfolding of the catalytic core domain PmScsCΔN, shows that its reduced state (T_m^{red} = 358 K) is more stable than its oxidised (T_m^{ox} = 345 K) form. Data are shown as the mean ± s.d., of three replicate experiments.* ***D.*** *Rate of dithiol oxidase catalytic activity of 80 nM PmScsC, PmScsCΔN and EcDsbC normalised to the activity of the control enzyme EcDsbA. Glutathione was used as the electron donor in the reaction. PmScsCΔN has an activity similar to that of the archetypal dithiol oxidase,*

EcDsbA, whereas trimeric PmScsC, the PmScsC RHP mutant and dimeric EcDsbC have negligible activity. Data are shown as the mean ± s.d., of three replicate experiments.

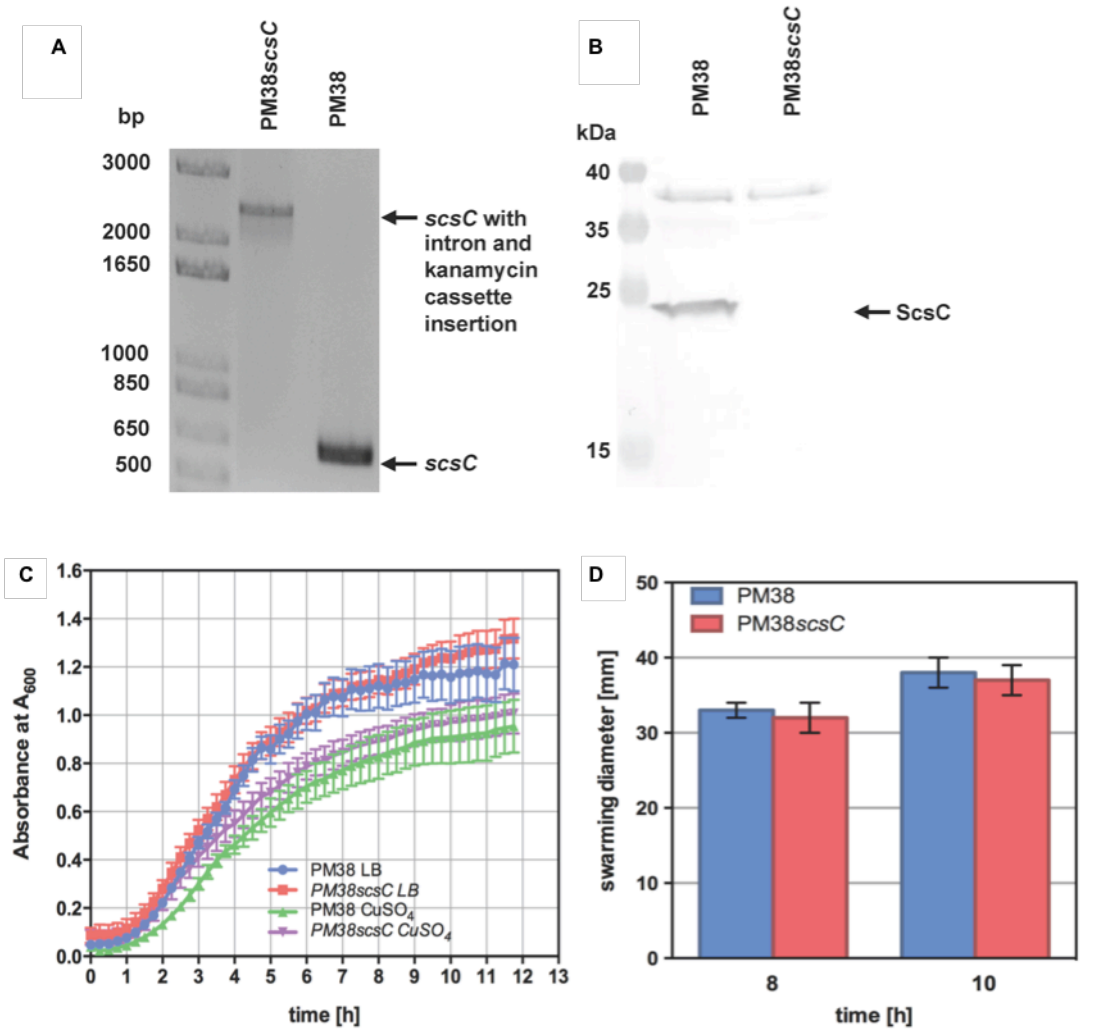


Supplementary Figure 2.2: Genotypic and phenotypic analysis of PM54 and PM54scsC. A. *Agarose gel electrophoresis of* scsC *PCR fragments obtained from PM54 and PM54*scsC*. An amplicon of increased size was obtained from PM54*scsC *(compared to PM54) using primers 5914 and 5915, confirming the correct insertion of the intron and kanamycin cassette in the* scsC *gene.* ***B.*** *Western blot analysis of whole cell lysates prepared from PM54, PM54*scsC*, PM54*scsC*(pSU2718), PM54*scsC*(pScsC), PM54*scsC*(pScsC AXXC) and PM54*scsC*(pScsCΔN) using an ScsC-specific polyclonal antiserum. Expression of wild-type ScsC and mutant forms (ScsC AXXC and ScsCΔN) was observed; no expression of ScsC was detected in PM54*scsC*.* ***C.*** *Growth analysis of PM54 and PM54*scsC *in LB broth at 37°C in the presence or absence of 1.5 mM CuSO4. There was no significant difference in the growth rate of PM54 (blue) and PM54*scsC *(red) in LB broth, or PM54 (green) and PM54*scsC *(purple) in LB broth containing 1.5 mM* $CuSO_4$*. Growth was monitored by measuring the optical density at 600nm (*OD_{600nm}*). Data*

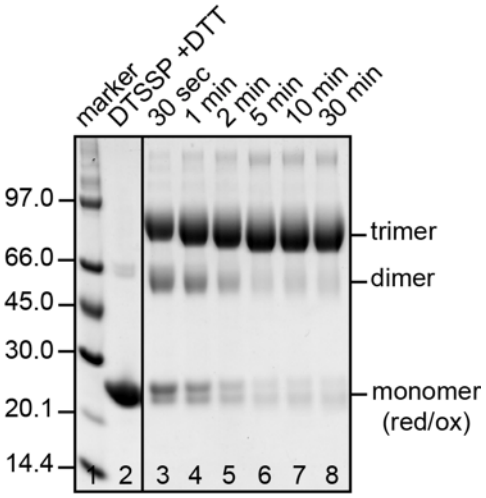
are shown as the mean ± s.d., of three replicate experiments. ***D.*** *Analysis of PM54 (black), PM54*scsC *(grey), PM54*scsC*(pSU2718) (light blue), PM54*scsC*(pScsC) (red), PM54*scsC*(pScsC AXXC) (dark blue) and PM54*scsC*(pScsCΔN) (green) swarming on LB agar at 37°C at 6 hours and 8 hours post-inoculation. There was no significant difference in the swarming of all strains examined on LB agar in the absence of added* $CuSO_4$*. Data are shown as the mean ± s.d. deviation, of three replicate experiments.*



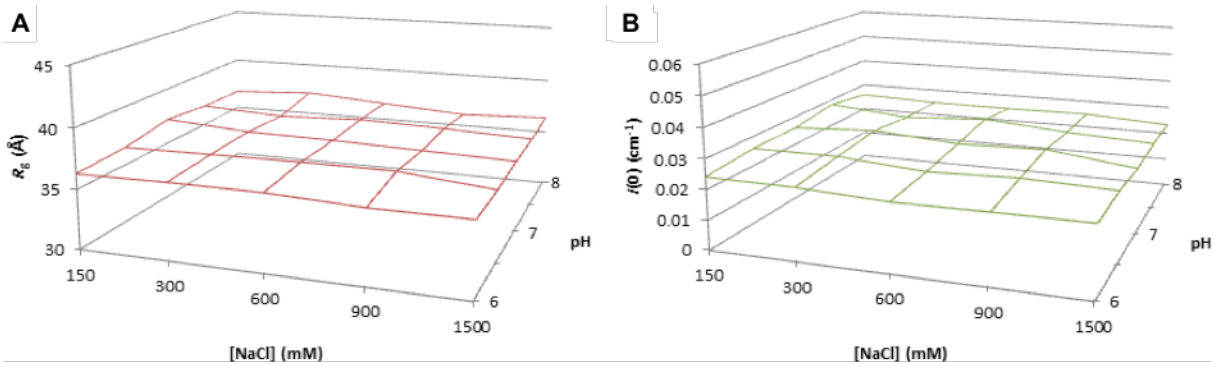
Supplementary Figure 2.3: Mutant phenotype and function of PmScsC. *Difference in swarming motility between wild type PM38 and PM38*scsC *in the presence of 1.5 mM CuSO4 on LB agar plates after incubation for 14 h at 37 °C. PM38*scsC *swarming motility (blue) was significantly reduced compared to wild type PM38 (black; (*P < *0.0001 for slope calculated by F-test). Data are shown as the mean ± s.d. deviation, of three replicate experiments (in some cases the error bars are too small to be visible).*



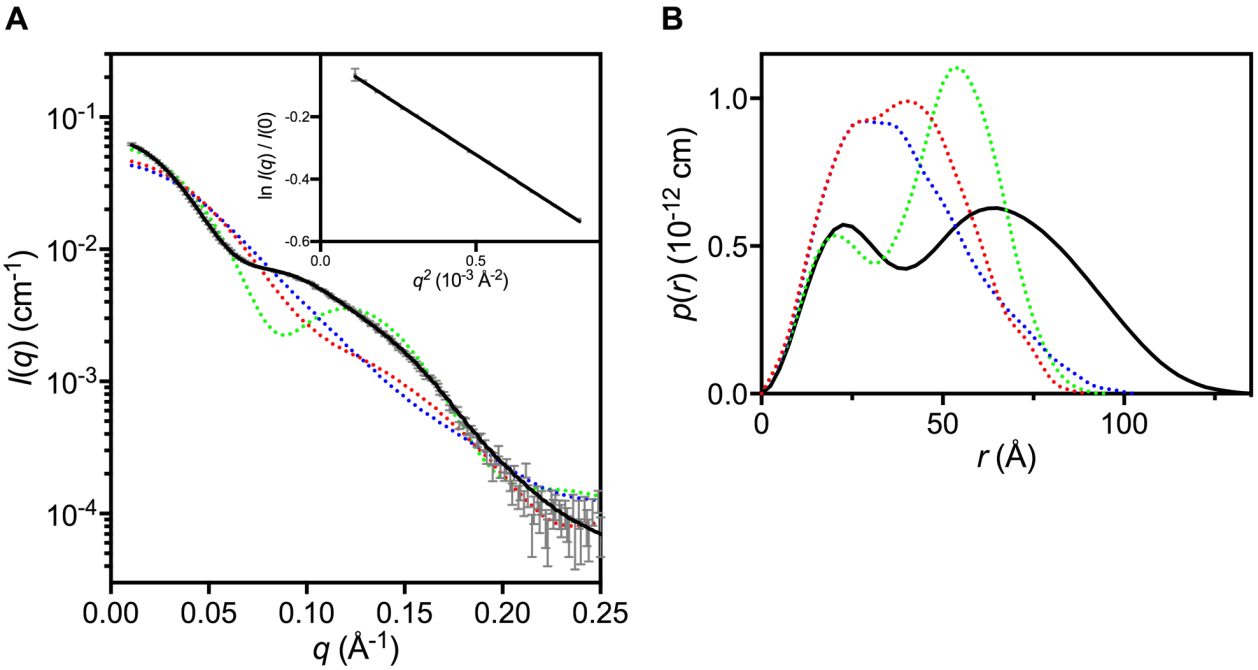
Supplementary Figure 2.4: Genotypic and phenotypic analysis of PM38 and PM38scsC. A. *Agarose gel electrophoresis of* scsC *PCR fragments obtained from PM38 and PM38*scsC*. An amplicon of increased size was obtained from PM38*scsC *(compared to PM38) using primers 5914 and 5915, confirming the correct insertion of the intron and kanamycin cassette in the* scsC *gene.* ***B****. Western blot analysis of whole cell lysates prepared from PM38 and PM38*scsC *using an ScsC-specific polyclonal antiserum. ScsC is expressed by PM38 but not by the PM38*scsC *mutant.* ***C.*** *Growth analysis of PM38 and PM38*scsC *in LB broth at 37°C in the presence or absence of 1.5 mM* $CuSO_4$*. There was no significant difference in the growth rate of PM38 (blue) and PM38*scsC *(red) in LB broth, or PM38 (green) and PM38*scsC *(purple) in LB broth containing 1.5 mM CuSO4. Growth was monitored by measuring the* OD_{600nm}*. Data are shown as the mean ± s.d., of three replicate experiments.* ***D.*** *Analysis of PM38 (blue) and PM38*scsC *(red) swarming on LB agar at 37°C at 8 hours and 10 hours post-inoculation. There was no significant difference in the swarming of PM38 and PM38*scsC *on LB agar in the absence of added* $CuSO_4$*. Data are shown as the mean ± s.d., of three replicate experiments.*



Supplementary Figure 2.5: SDS gel of cross-linked PmScsC. *Denaturing SDS-PAGE (-DTT) of purified PmScsC, incubated for 30 s – 30 min with a disulfide-bonded chemical cross-linker DTSSP. Within 30 s a trimer is evident (lanes 3-8) that can be disrupted by DTT added to the sample after 30 min time (DTSSP + DTT).*

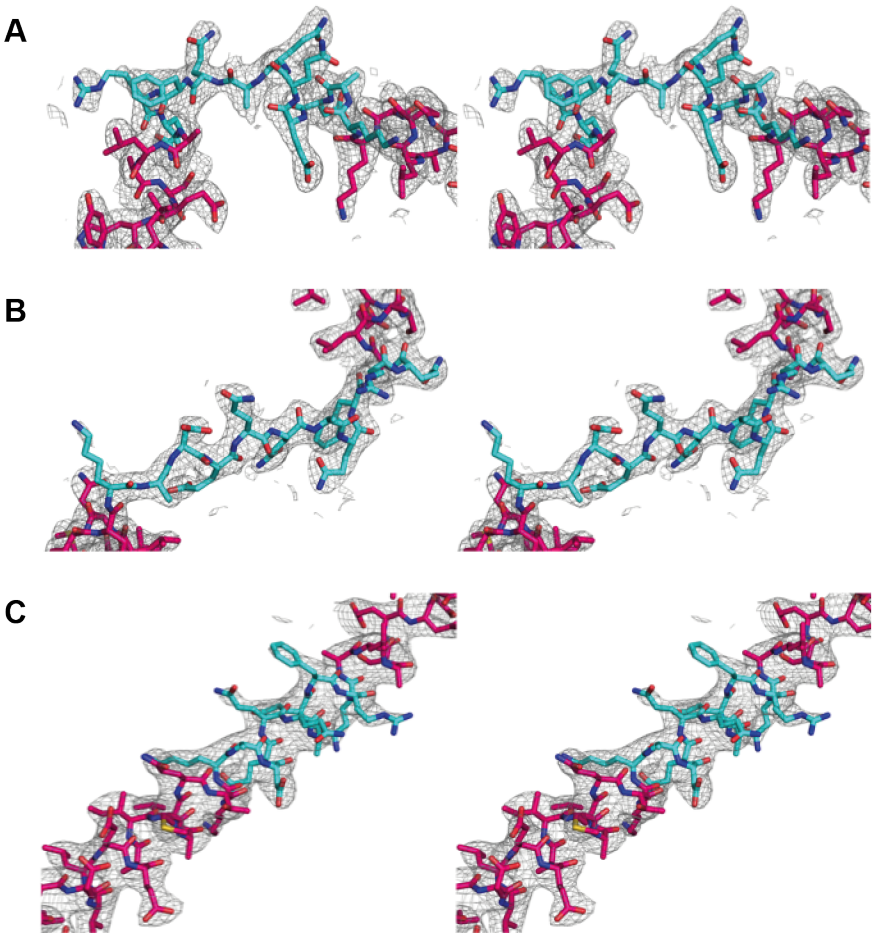


Supplementary Figure 2.6: PmScsC NaCl/pH gradient scattering data. A. *The variation of the radius of gyration is shown as a function of both pH and NaCl concentration. There is no significant or systematic variation with NaCl or pH, indicating that the conformation/flexibility of the protein is not dependent on solution conditions, but is an inherent property of the protein.* ***B.*** *The variation in the forward scattering,* I*(0), shown as a function of both pH and NaCl concentration. The values are corrected for the change in contrast of the protein in the varying concentration of NaCl. The plot shows the oligomeric state of the protein is independent of both NaCl concentration and pH.*



Supplementary Figure 2.7: PmScsC RHP mutant scattering data. A. *Small-angle X-ray scattering data collected from the PmScsC RHP mutant (grey) and the calculated scattering profile of the ensemble model overlayed in black (SASBDB: SASDBW6). The predicted scattering profile of each of the crystal structures is also shown (dashed lines: PDB: 4XVW compact, red; PDB: 5IDR transitional, blue; PDB: 5ID4 extended, green). The agreement between the experimental data and an ensemble model is excellent, yielding* χ^2*=1.8 (compared to* χ^2*=1708.1 (compact);* χ^2*=1710.3 (transitional);* χ^2*=1365.5). The Guinier region (inset) of the scattering data is linear, consistent with a monodisperse solution.* ***B.*** *Pair distance distribution function derived from the scattering data, showing the maximum dimension of the particles in solution is 135 Å. Also shown is the calculated* p(r) *for each of the other structures (dashed lines: compact, red; transitional, blue; extended, green), showing a maximum dimension of 90, 105, and 100 Å, respectively. The* p(r) *curves from each of the crystal structures are markedly different to the experimentally determined* p(r) *curve.*

Supplementary Figure 2.8: Sequence alignment of PmScsC and StScsC. *Secondary structure elements are shown for PmScsC (above the amino acid sequence) and StScsC (below the amino acid sequence). Secondary structure elements for PmScsC are coloured using the colour scheme from Figure 2.2A/B, whereas those for StScsC are shown in orange. The CXXC motif is shaded in yellow, and the* cis*Pro motif found in most thioredoxin fold proteins is highlighted in grey.*



Supplementary Figure 2.9: Stereo images of representative electron density in the PmScsC crystal structures. *The electron density of the flexible linker region in* ***A.*** *compact PmScsC (PDB: 4XVW)* ***B.*** *the intermediate protomer of the transitional structure (PDB: 5IDR) and* ***C.*** *extended PmScsC (PDB: 5ID4). In each case the linker adopts a different secondary structure, forming* ***A.*** *loop,* ***B.*** *β-strand or* ***C.*** *α-helix. The residues 39-KADEQQAQFRQ-49 are shown in cyan in each case and residues that form the trimerisation stem are coloured magenta. The likelihood-weighted* $2F_o - F_c$ *electron density maps (contoured to 1σ) are represented by the wire mesh. The maps were generated using phenix.maps and missing* F_o *were filled with* F_c*.*

2.10.3 Supplementary Notes

Supplementary Note 2.1: The codon optimised gene sequence used for expression of PmScsC(22-243) is:

```
GCTGCTCTGAATGCTGCCCAAGAAAAAGAAGTGCGTGCACTGGTTCGCGACACCCTGGTCAGCAAC
CCGGAAATTCTGGAAGAAGCAATCATGGCTCTGCAGACGAAAAAAGCGGATGAACAGCAAGCCCAG
TTTCGTCAAGCACTGGCTAGCGAACATGATGCCCTGTATAACGACGCAGCATCTCCGCGCATTGGT
GCAAAAGATGCCAAACTGGTGCTGGTTTCTTTTACCGACTATAATTGCCCGTACTGTAAACGTTTC
GATCCGCTGCTGGAAAAAATCACGGAACAGTACCCGGACGTCGCAGTGATTATCAAACCGCTGCCG
TTTAAAGGTGAAAGCTCTGCGAAAGCCTCACAGGCTGTTCTGTCGGTCTGGAAAGAAGATCCGAAA
GCATTCCTGGCTCTGCACCAGCGTCTGATGCAGAAGAAAACCATGCTGGATAACGCAAGCATTGAA
GACGCTATGAAAAGTACCAATACGTCCAAAATCAAACTGACCGATGACAGTCTGAAAACGCTGCAG
AACAATCTGGAACTGTCCCGCAAACTGGGCATTCAAGGTACCCCGGCGACGGTTATTGGCGATACC
ATCCTGCCGGGTGCCGTGGACTATGACCAACTGGAAATTATCGTGAAAGAACAACTGGCAAAAGTG
AAAAAATGA
```

2.10.4 Supplementary Movies

Supplementary Movie 2.1: Conformation of the flexible linker throughout morph between PmScsC crystal structures.

https://media.nature.com/original/nature-assets/ncomms/2017/170719/ncomms16065/extref/ncomms16065-s2.mov

Supplementary Movie 2.2: Morph between EcDsbC crystal structures.

https://media.nature.com/original/nature-assets/ncomms/2017/170719/ncomms16065/extref/ncomms16065-s3.mov

Supplementary Movie 2.3: Morph between human PDI crystal structures.

https://media.nature.com/original/nature-assets/ncomms/2017/170719/ncomms16065/extref/ncomms16065-s4.mov

Supplementary Movie 2.4: Morph between PmScsC crystal structures.

https://media.nature.com/original/nature-assets/ncomms/2017/170719/ncomms16065/extref/ncomms16065-s5.mov

Chapter 3

3. Characterisation of engineered PmScsC variants

3.1 Chapter foreword

This chapter is a reproduction of a research article that has been submitted to *Acta Crystallographica Section D*. This manuscript focuses on two engineered PmScsC variants, PmScsCΔN and PmScsCΔLinker. The crystal structure of PmScsCΔN is reported along with the SAXS analysis of this variant. PmScsCΔLinker is functionally and structurally characterised in this paper. Changes have been made to the submitted manuscript to conform to the thesis submission requirements; these include the overall formatting and numbering of the figures, tables and references.

3.1.1 Reference

Furlong, E. J., Kurth, F., Premkumar, L., Whitten, A. E., Martin, J. L. (2018) Engineered variants provide new insight into the structural properties important for activity of the highly dynamic, trimeric protein disulfide isomerase, PmScsC. Available on bioRxiv; doi: 10.1101/421420. Submitted to *Acta Crystallographica Section D*.

3.1.2 Contribution

I significantly contributed to the conception and design of the project, the analysis and interpretation of the research data and the drafting of this submitted manuscript. I designed the PmScsCΔLinker variant and created the expression construct for this protein. I expressed and purified PmScsCΔLinker and assessed the protein for disulfide isomerase and dithiol oxidase activity. I crystallised PmScsCΔLinker, collected and processed diffraction data and solved and refined the crystal structure of this protein. I analysed both PmScsCΔN and PmScsCΔLinker crystal structures and my interpretation of these structures is presented in the manuscript. I contributed to the collection of the SAXS data reported in this manuscript. I also wrote the first draft of the article and designed and created the figures.

Engineered variants provide new insight into the structural properties important for activity of the highly dynamic, trimeric protein disulfide isomerase, PmScsC

Emily J. Furlong[a,b], Fabian Kurth[a], Lakshmanane Premkumar[a], Andrew E. Whitten[c*] and Jennifer L. Martin[a,b*]

Affiliations

[a]Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD 4072, Australia. [b]Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD 4111 Australia. [c]Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW 2234, Australia.

Currently: FK Bristol-Myers Squibb, Arnulfstraße 29, 80636 Munich, Germany; LP Department of Microbiology and Immunology, School of Medicine, University of North Carolina, Chapel Hill, NC 27514 U.S.A

*Correspondence

AEW awh@ansto.gov.au and JLM jlm@griffith.edu.au

Publication timeline

Submitted to BioRxiv: 19 September 2018

Submitted to *Acta Crystallographica Section D*: 29 September 2018

3.2 Abstract

Suppressor of copper sensitivity protein C from *Proteus mirabilis* (PmScsC) is a homotrimeric disulfide isomerase that plays a role in copper tolerance – a key virulence trait of the uropathogen. Each protomer of the enzyme has an N-terminal trimerisation stem (59 residues) containing a flexible linker (11 residues) connected to a thioredoxin-fold-containing catalytic domain (163 residues). Here, we characterise two PmScsC variants, PmScsCΔN and PmScsCΔLinker. PmScsCΔN, is an N-terminally truncated form of the protomer with two helices of the trimerisation stem removed, generating a protein with dithiol oxidase rather than disulfide isomerase activity. The crystal structure of PmScsCΔN reported here reveals – as expected – a monomer that is structurally similar to the catalytic domain of native PmScsC. The second variant PmScsCΔLinker was designed to remove the 11 amino acid linker and we show that it generates a protein that has neither disulfide isomerase nor dithiol oxidase activity. The crystal structure of PmScsCΔLinker reveals a trimeric arrangement, with the catalytic domains packed together very closely. Small angle X-ray scattering analysis found that native PmScsC is predominantly trimeric in solution even at low concentration, whereas PmScsCΔLinker exists as an equilibrium between monomeric, dimeric and trimeric states, with the monomeric form dominating at low concentrations. These findings increase our understanding of disulfide isomerase activity, showing how (i) oligomerisation, (ii) spacing between, and (iii) dynamic motion of, catalytic domains in PmScsC all contribute to its native function.

3.3 Introduction

Protein disulfide isomerases are enzymes that proofread and shuffle incorrect disulfide bonds in misfolded protein substrates, and are important for the correct folding and function of many secreted proteins (17,173). The prototypical bacterial disulfide isomerase is Disulfide bond (Dsb) protein C from *Escherichia coli* (EcDsbC) (4,5,42). EcDsbC functions as a dimer; when the N-terminal dimerisation domain is deleted the resulting protein lacks disulfide isomerase activity (41) and it is also unable to interact with its redox partner EcDsbD (44). Each EcDsbC protomer has a thioredoxin-fold catalytic domain with a redox active motif consisting of two cysteines separated by two other amino acids (CXXC) (39). The disulfide isomerase activity of EcDsbC requires that the catalytic cysteines are in the dithiol reduced form (174) and this form is generated by interaction of EcDsbC with its redox partner membrane protein EcDsbD (44).

Recently, we reported the structural and functional characterisation of *Proteus mirabilis* suppressor of copper sensitivity protein C (PmScsC) (175). Like DsbC, PmScsC is a protein disulfide isomerase with a thioredoxin-fold catalytic domain (175), redox-active cysteines (175), and a redox partner membrane protein (PmScsB) that reduces the active site cysteines (176). Unlike DsbC, PmScsC is trimeric rather than dimeric (175). Three crystal structures and SAXS analysis also revealed that PmScsC is highly dynamic (175). The region responsible for this conformational flexibility is an 11 amino acid linker within the trimerisation stem of the protein. Replacement of the flexible linker with a rigid helical peptide linker (PmScsC RHP), limits the rigid body motion of the catalytic domain relative to the trimer stalk, and generates a protein that is inactive as a disulfide isomerase. Moreover, PmScsC lacking the trimerisation stem (N-terminal 41 residues, PmScsCΔN) is also inactive as a disulfide isomerase but does have dithiol oxidase activity (175). Neither the native protein, nor the PmScsC RHP variant, exhibited dithiol oxidase activity. The difference in activity profile between the native protein (disulfide isomerase) and PmScsCΔN (dithiol oxidase) was proposed to be a consequence of their different oligomeric states.

Here we report the structural characterisation of PmScsCΔN and the structural and functional characterisation of PmScsCΔLinker (a variant in which the 11-amino acid flexible linker is deleted, Figure 3.1). The crystal structure of PmScsCΔN reveals – as expected – a monomeric protein that closely resembles the catalytic domain from the full-length PmScsC protomer. We show that the PmScsCΔLinker variant lacks both disulfide

isomerase and dithiol oxidase activity. The crystal structure reveals a trimeric arrangement for PmScsCΔLinker with the catalytic domains closely packed against each other. Conversely, SAXS analysis indicated that the oligomeric state of PmScsCΔLinker is concentration-dependent, with a monomeric form present at low concentrations and a dimeric form dominating at higher concentrations. For comparison, we analysed a concentration series of full-length native PmScsC by SAXS showing that the dominant species of the native protein in solution is trimeric, even at very low concentrations.

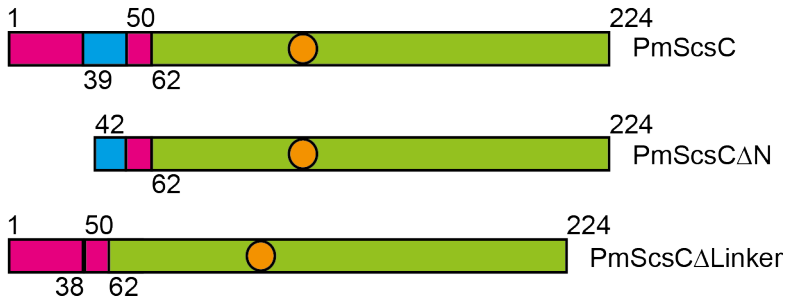


Figure 3.1: Schematic representation of the PmScsC variants used in this study. *The region corresponding to the catalytic domain is shown in green, the trimerisation stem is coloured magenta and the region representing the flexible linker is shown in cyan. The orange circle represents the approximate position of the redox active cysteines.*

3.4 Materials and methods

3.4.1 Molecular biology, protein expression and purification

The constructs for the expression of PmScsC and PmScsCΔN were created previously (175) using the UniProt sequence B4EV21, without the predicted periplasmic signaling sequence (i.e. residues 22-243). To make the PmScsCΔLinker expression construct, the DNA encoding the flexible peptide linker (residues 39-KADEQQAQFRQ-49) in PmScsC was deleted from the *P. mirabilis scsC* gene using overlap extension PCR, with the primers 5'-GCAATCATGGCTCTGCAGACGAAAGCACTGGCTAGCGAACATGATGCC and 5'-GGCATCATGTTCGCTAGCCAGTGCTTTCGTCTGCAGAGCCATGATTGC. Ligation independent cloning was then used to insert the mutated gene into pMCSG7. Sequencing of the construct confirmed that the encoded protein was in-frame with an N-terminal His_6-tag and TEV protease cleavage site and revealed that residue 6 of PmScsCΔLinker was also mutated from asparagine to lysine. The protein residues are numbered based on the sequence of full-length PmScsC, after TEV protease cleavage (starting $S_1N_2A_3$...). All proteins were expressed in *E. coli* and purified using immobilised metal affinity, TEV cleavage and size exclusion chromatography as previously described (175). SDS-PAGE with Coomassie stain was used to estimate that the purity of each protein was >95%. The concentration of the proteins was determined by measuring the A_{280} of the samples and adjusting for theoretical A_{280} of a 1 mg/ml solution of the protein (Abs 0.1%). The theoretical Abs 0.1% values determined from Protparam (140) were 0.528, 0.643 and 0.558 for PmScsC, PmScsCΔN and PmScsCΔLinker, respectively.

3.4.2 Disulfide isomerase and dithiol oxidase assays

The disulfide isomerase and dithiol oxidase activity of PmScsCΔLinker was determined as described previously (175). Briefly, to determine the disulfide isomerase activity, the ability of the protein to refold scrambled RNase A (scRNase A) was assessed. 10 μM PmScsCΔLinker was incubated with 40 μM scRNase A and at various time points, samples were taken and the RNAse A activity was measured in a spectrophotometric assay with a final concentration of 3 mM cytidine 3',-5'-cyclic monophosphate (cCMP). Results are reported as a percentage of the activity of the RNase only control and are adjusted for the scrambled RNase A only negative control. The dithiol oxidase activity of PmScsCΔLinker was assessed by measuring the protein's ability to form a disulfide bond between the two cysteines in a synthetic peptide, over time. The peptide substrate has 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) bound to europium at the N-terminus and lysine methoxy-coumarin amide (MCA) at the C-terminus. As the

cysteines become oxidised, the termini of the peptide are brought within close proximity, which causes an increase in fluorescence. PmScsCΔLinker (80 nM) was incubated with 8 μM of the peptide substrate and the increase in fluorescence over time was monitored using a Synergy H1 Hybrid plate reader (BioTek, Vermont, USA) reading fluorescence at excitation and emission wavelengths of 340 nm and 615 nm, respectively. The rate of peptide oxidation is reported relative to the activity of the prototypical dithiol oxidase EcDsbA. The disulfide isomerase and dithiol oxidase assays were repeated two and three times, respectively and the mean and standard deviation for each is reported.

3.4.3 Crystallisation

Initial crystallisation hits for both PmScsCΔN and PmScsCΔLinker were identified in 96-well hanging-drop vapour-diffusion experiments using commercial crystallisation screens, set up with a Mosquito robot (TTP LabTech Ltd, Melbourn, UK) at the University of Queensland Remote Operation Crystallisation and X-ray (UQ ROCX) Diffraction Facility. All crystallisation experiments were incubated at 293 K and monitored using a Rock Imager with Rock Maker software (Formulatrix Inc., Waltham, Massachusetts, USA).

Well diffracting, rod-shaped crystals of PmScsCΔN were obtained in a 96-well screening plate in a drop containing 200 nL of the PEGRx™ HT (Hampton Research, Aliso Viejo, California, USA) C10 condition, 0.1 M Sodium acetate trihydrate pH 4.5, 30% w/v PEG-mono methylether 5000, and 200 nL of 38 mg/ml PmScsCΔN in 10 mM HEPES pH 7.4. Perfluoropolyether (Hampton Research, Aliso Viejo, California, USA) was used to cryoprotect the crystals before they were flash frozen in liquid nitrogen.

Needle-like crystals of PmScsCΔLinker were obtained in well F12 of the commercial screen JCSG-*plus*™ (Molecular Dimensions Limited, New-Market, UK) and then optimised in a 24-well plate grid screen around the original condition. This optimisation resulted in more robust and well diffracting crystals. Data were collected from a crystal grown in 1 μL of 0.1 M HEPES pH 7, 32% v/v Jeffamine M-600 and 1 μL of 40 mg/mL PmScsCΔLinker in 10 mM HEPES pH 7.4, 150 mM NaCl. The crystal was cryoprotected in the crystallisation condition plus 20% ethylene glycol before being flash cooled in a cryostream (100 K).

3.4.4 Data collection, structure solution and refinement

Data for the PmScsCΔN structure were collected using a wavelength of 0.95370 Å at 100 K on the Australian Synchrotron MX2 beamline using the BluIce software (150). The data were processed and scaled with the trigonal space group P312 using XDS (151), Pointless (177) and SCALA (177). Phases were obtained by molecular replacement in Phaser (155), using *Salmonella enterica* serovar Typhimurium ScsC (PDB: 4GXZ) as the model. The initial model was improved by iterative model building in Coot (178) and refinement in PHENIX (154). The quality of the final PmScsCΔN model was assessed using MolProbity (164) and had an R_{free} value of 21.5% and R_{work} value of 19.6%, at a resolution of 2.15 Å. The coordinates and structure factors are deposited to the Protein Data Bank (179,180) under the code 4YX8.

Data for the PmScsCΔLinker structure were collected at UQ ROCX under cryogenic conditions (100 K) using a Rigaku FR-E Superbright X-ray generator, producing X-rays at a wavelength of 1.54187 Å, and a Rigaku Saturn 944 CCD area detector. The data collection was performed with the Rigaku CrystalClear (v 2.0) program. Data were processed in iMosflm (v 7.1.3) (181) with the space group H32 (also known as R32:h) and scaled in Aimless (152) as implemented in the CCP4 suite (v 6.5.008) (160). Due to detector distance constraints at UQ ROCX, diffraction data were limited to a highest resolution of 2.08 Å. The crystals diffracted strongly, resulting in very high mean <I/σI> values (27.2 overall and 11.8 in the highest resolution shell). The structure was phased by molecular replacement using Phaser (155) and the catalytic domain of the compact PmScsC crystal structure (residues 47-224, PDB: 4XVW) was used as the search model. The trimerisation domain (without the 11 amino acid flexible linker) was then manually built into the electron density in Coot (178). The structure was refined using rounds of phenix.refine and manual adjustment in Coot, with reference to MolProbity (164). The final R_{free} and R_{work} values are 20.96% and 17.42% respectively, at a resolution of 2.08 Å. The coordinates and structure factors are deposited with the Protein Data Bank (179,180) under the code 6MHH. All structural figures were made in PyMOL (182) and structural alignments were performed in Coot using the superpose command (178) which aligns Cαs by least squares. The Contact program from the CCP4 suite (160) was used to analyse the interactions between the trimerisation stems of PmScsC/PmScsCΔLinker.

3.4.5 SAXS

SAXS data were collected on the SAXS-WAXS beamline at the Australian Synchrotron (142). Data reduction was carried out using Scatterbrain software (v 2.71) (143), and corrected for solvent scattering, sample transmission, and detector sensitivity. The 0.1, 0.5, 2.0, 5.0 and 10.0 mg/mL samples of PmScsCΔN were made by diluting a 30 mg/mL stock in 10 mM HEPES pH 7.4, 150 mM NaCl. For PmScsC and PmScsCΔLinker, samples were taken throughout centrifugal concentration and the concentration of these samples was adjusted to 0.1, 0.5, 2.0, 5.0 or 10.0 mg/mL in 10 mM HEPES pH 7.4, 150 mM NaCl. These samples, along with the blank (10 mM HEPES pH 7.4, 150 mM NaCl) were loaded into a 96-well plate (Table 2) (183). The estimated molecular mass was calculated from the Porod volume (184) and from $I(0)$ using contrast and partial specific volumes determined from the protein sequences using MULCh (v 1.1) (145). For PmScsCΔN, data processing and Guinier analysis was performed using Primus (v 3.2) (185). The pair-distance distribution function ($p(r)$) was generated from the experimental data using *GNOM* (v 4.6) (146), from which $I(0)$, R_g and D_{max} were determined. The program *DAMMIN* (v 5.3) (186) was used to generate 16 dummy-atom models for the protein, which were averaged using the program *DAMAVER* (v 2.8.0) (187), and the resolution of the averaged structures was estimated using SASRES (188). All 16 dummy-atom models were used in the averaging procedure. The model scattering curve for PmScsCΔN (SASBDB ID: SASDEK4) was calculated from the crystal structure (PDB ID: 4YX8) using CRYSOL (v 2.8.3) (189).

For PmScsC (SASBDB ID: SASDER4) and PmScsCΔLinker (SASBDB ID: SASDEQ4), data were modelled as a linear combination of model scattering curves to quantify the proportion of different oligomeric states in solution (190). The structure of PmScsC monomer (chain ID: A), dimer (chain ID: A, B) and trimer (chain ID: A, B, C) were taken from a CORAL model previously optimised against SAXS data for PmScsC (SASDB94) (175). The structure of PmScsCΔLinker monomer (chain ID: A), dimer (chain ID: A, B) and trimer (chain ID: A, B, C) were taken from the crystal structure of PmScsCΔLinker (PDB ID: 6MHH). Scattering curves for each oligomer were calculated using CRYSOL (v 2.8.3) (189). The calculated scattering curves (with units of e^2) were multiplied by r_e^2 ($cm^2\ e^{-2}$) × N_A (mol^{-1}) × 10^{-3} (L cm^{-3}) = 4.78181 × 10^{-5} ($cm^{-1}\ e^{-2}\ mol^{-1}$ L). This normalisation permits the coefficients arising from a fit to experimental data on an absolute scale to be interpreted as the molar concentration of each oligomer present in solution. For PmScsC, a linear combination of monomer and trimer scattering curves were fit to the scattering data at

each concentration. To reduce the number of free parameters, the concentration of monomer for each SAXS curve was optimised together with a common equilibrium constant, K_t (3Monomer $\overset{K_t}{\leftrightarrow}$ Trimer), and the concentration of trimer was calculated from K_t and the momomer concentration. For PmScsCΔLinker, a linear combination of monomer, dimer and trimer scattering curves were fit to the scattering data at each concentration. To reduce the number of free parameters, the concentration of monomer for each SAXS curve was optimised together with two common equilibrium constants, K_d (2Monomer $\overset{K_d}{\leftrightarrow}$ Dimer), and K_t (Monomer+Dimer $\overset{K_t}{\leftrightarrow}$ Trimer). The concentration of dimer was calculated from K_d and the monomer concentration, while the concentration of trimer was calculated from K_t and the concentration of monomer and dimer.

3.5 Results

3.5.1 PmScsCΔN crystal structure and SAXS analysis

The design and biochemical characterisation of PmScsCΔN has been reported previously (175). The present work reports the crystal structure of this variant. Crystals of the PmScsCΔN mutant diffracted to a resolution of 2.15 Å at the Australian Synchrotron. The crystal structure of the mutant was solved using molecular replacement and there was only one PmScsCΔN molecule in the asymmetric unit (Table 3.1, Figure 3.2A, B). In the crystal packing, PmScsCΔN makes contact with 6 symmetry related molecules, but none involve packing against the catalytic motif or the positively charged surface patch surrounding it. The structure of PmScsCΔN is very similar to the structure of the catalytic domain present in each of the three previously reported PmScsC structures (Figure 3.2C, RMSD: 0.51-0.68 Å, Cα atoms of 174 residues aligned, between chain A of each structure). Small-angle scattering (Table 3.2) was used to determine the low-resolution solution structure of PmScsCΔN (Figure 3.3). The scattering data for PmScsCΔN is representative of a monomeric globular particle and is consistent with the scattering curve calculated from the crystal structure of PmScsCΔN (PDB ID: 4YX8). The crystal structure also shows good agreement with the dummy atom model derived from the scattering data (Figure 3.3C).

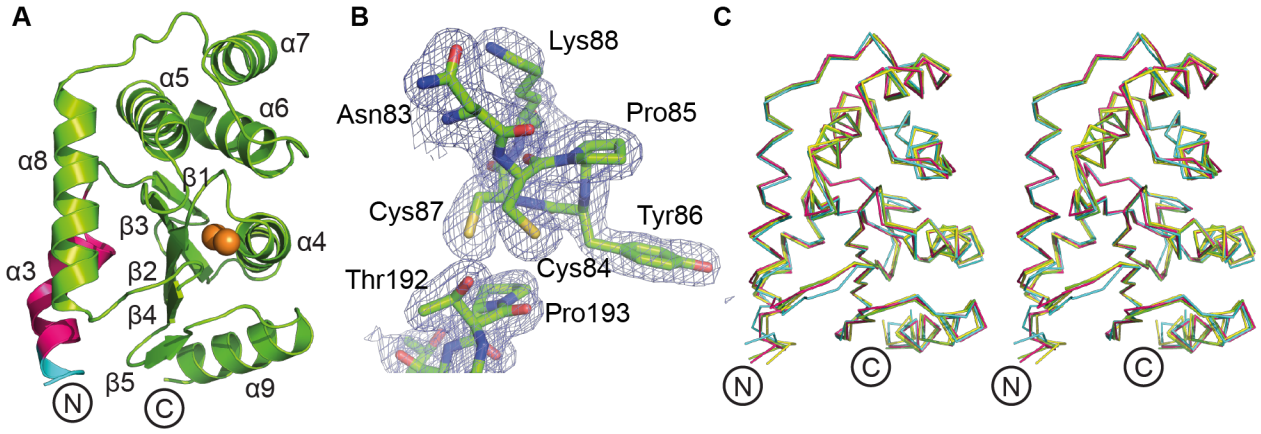


Figure 3.2: Structural characterisation of PmScsCΔN. A. *Structure of the PmScsCΔN variant. Structural features are coloured as in Figure 3.1 and the secondary structure elements are labelled based on the native PmScsC structures (175).* ***B.*** *Electron density of the catalytic region in PmScsCΔN. The wire mesh represents the $2F_o$-F_c electron density map (generated in phenix.maps) contoured to 1σ and shown within a 2 Å radius of each atom.* ***C.*** *The structural alignment of PmScsCΔN (green) with residues 46-221 of chain A from the compact (cyan, PDB ID: 4XVW), transitional (yellow, PDB ID: 5IDR) and extended (magenta, PDB ID: 5ID4) PmScsC structures (175) (RMSD: 0.51-0.68 Å, 174 residues aligned). This superimposition shows that the structures are similar whether or not the trimerisation domain is present.*

Table 3.1: Crystallography statistics for PmScsCΔN and PmScsCΔLinker.

Data collection	**PmScsCΔN PDB: 4YX8**	**PmScsCΔLinker PDB: 6MHH**
Wavelength (Å)	0.9537	1.541870
Resolution (Å)	45.68-2.15 (2.22-2.15)[a]	24.99-2.08 (2.14-2.08)
Space group	P312	H32 (R32:h)
Unit cell dimensions		
a, *b*, *c* (Å)	105.5, 105.5, 35.13	64.04, 64.04, 299.83
α, β, γ (º)	90, 90, 120	90, 90, 120
Number measured reflections	265709	148450
Number of unique reflections	12203	14715 (1135)
R_{merge}	0.13 (0.861)	0.059 (0.140)
Mean <I/σI>	21.6 (4.9)	27.2 (11.8)
Redundancy	21.9 (20.5)	10.1 (7.8)
Completeness (%)	99.7 (97.0)	99.9 (100.0)
Wilson-B	28.30	21.85
Refinement		
Number of monomers in a/u	1	1
Resolution used in refinement (Å)	35.13-2.15 (2.22-2.15)	24.99-2.08
Number of unique reflections	12197	14704
R_{free} (%)	21.5	20.96
R_{work} (%)	19.6	17.42
Number of protein atoms (non-H)	1326	1671
Number of ligand atoms	0	0
Number of waters	80	131
B factors (Å^2)		
Average	47.8	30.5
Protein atoms	47.5	30.3
Ligands	NA	NA
Waters	53.6	33.4
RMSD Bond length (Å)	0.008	0.003
RMSD Bond angles (º)	1.065	0.687

MolProbity results		
Ramachandran favoured (%)	98	99.5
Ramachandran outlier (%)	0.6	0.0
Clashscore	1.13	1.17

[a]Values in parentheses refer to the highest resolution shell.

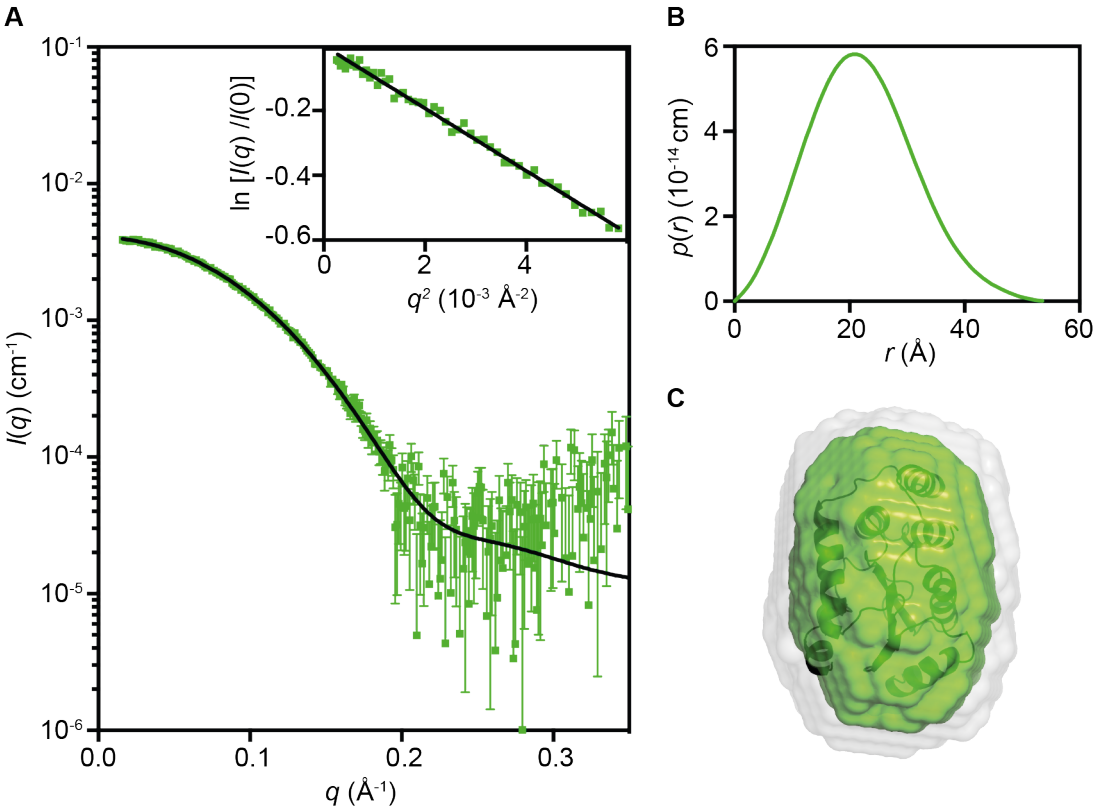


Figure 3.3: Small angle X-ray scattering data for PmScsCΔN. A. *Measured scattering data for PmScsCΔN. The scattering profile of rigid-body models are shown as solid black lines overlaid on the scattering data for PmScsCΔN (χ^2 = 1.12; CorMap test (148), 302 points, C=14, P=0.018). Inset: Guinier plot for PmScsCΔN (R^2 = 0.991)* ***B.*** *The pair-distance distribution function, p(r), derived from the scattering data are indicative of a globular structure for PmScsCΔN (multiplied by a factor of 4 for clarity) with a maximum dimension of ~55 Å.* ***C.*** *Probable shape of PmScsCΔN obtained from the filtered average of 16 dummy-atom models (green envelope): χ^2 = 1.079 ± 0.001; NSD = 0.522 ± 0.007; Resolution = 19 ± 2. Images in C were generated using PyMol, where the grey shapes represent the total volume encompassed by the aligned dummy-atom models and the corresponding rigid-body model is shown aligned to the filtered model.*

Table 3.2: SAXS data collection and analysis details.

Data Collection Parameters	**PmScsCΔN**	**PmScsC**	**PmScsCΔLinker**
Instrument	SAXS-WAXS (Australian Synchrotron)		
Beam geometry	Point		
Wavelength (Å)	1.033		
Sample to detector distance (m)	3.330		
q-range ($Å^{-1}$)	0.008 – 0.363		
Temperature (K)	283		
Absolute intensity calibration	Water		
Exposure time (s)	38 (38 × 1 s exposures)		
Configuration	Concentration series from 96-well plate		
Protein concentration range (mg/mL)	0.1 – 10.0		
Sample details			
Extinction coefficient (A_{280}, 0.1% *w/v*)	0.643	0.528	0.558
Partial specific volume (cm^3g^{-1})	0.742	0.741	0.743
Contrast, $\Delta\rho$ (10^{10} cm^{-2})	2.79	2.80	2.78
Molecular mass [from sequence] (kDa)	20.2	24.8	23.4
Protein concentration (mg/mL)	0.50*	0.1, 0.5, 2.0, 5.0, 10.0	0.1, 0.5, 2.0, 5.0, 10.0
Structural parameters			
$I(0)$ (cm^{-1}) [from Guinier]	0.004055 ± 0.00001	-	-
R_g (Å) [from Guinier]	17.0 ± 0.1	-	-
$I(0)$ (cm^{-1}) [from $p(r)$]	0.004053 ± 0.00001	-	-
R_g (Å) [from $p(r)$]	17.0 ± 0.1	-	-
D_{max} (Å)	55 ± 2	-	-
Porod volume ($Å^3$)	22200 ± 1000	-	-
Molecular mass determination			
Molecular mass [from $I(0)$] (kDa)	11 ± 2	-	-
Molecular mass [from Porod] (kDa)	19 ± 1	-	-

*PmScsCΔN: $I(0)$ = 0.00140 ± 0.00002 cm^{-1}, R_g = 16.9± 0.3 Å, M = 19.8 kDa (0.1 mg/mL); $I(0)$ = 0.004053 ± 0.000001 cm^{-1}, R_g = 17.0 ± 0.1 Å, M = 19.1 kDa (0.5 mg/mL); $I(0)$ = 0.01839 ± 0.00003 cm^{-1}, R_g = 17.1 ± 0.1 Å, M = 19.3 kDa (2.0 mg/mL); $I(0)$ = 0.05463 ± 0.00005 cm^{-1}, R_g = 17.2 ± 0.1 Å, M = 19.6 kDa (5.0 mg/mL); $I(0)$ = 0.10756 ± 0.00005

cm^{-1}, R_g = 17.0 ± 0.1 Å, M = 19.2 kDa (10.0 mg/mL). Masses from the Porod volume are used as concentration estimates yielding masses from $I(0)$ inconsistent with the expected mass. The 0.5 mg/mL was used for subsequent data analysis.

3.5.2 Activity of PmScsCΔLinker

A key structural feature of native PmScsC is an 11 amino acid flexible linker, situated in the trimerisation stem and linking to the catalytic domain. This short region facilitates the conformational variations observed in the crystal structures and the dynamic motion of the trimeric protein observed in SAXS analyses (175). To better understand the role of the flexible linker we designed a PmScsCΔLinker variant in which the 11-residue linker is deleted, and characterised its function in standard assays. We first assessed the activity of PmScsCΔLinker in the scrambled RNase A assay and found that it was poorly active, with baseline activity like that of a typical dithiol oxidase such as EcDsbA and PmScsCΔN (Figure 3.4A). We then assessed its activity in the model peptide dithiol oxidase assay, in which EcDsbA and PmScsCΔN are both active. However, unlike the monomeric proteins EcDsbA and PmScsCΔN, PmScsCΔLinker was also inactive in the dithiol oxidase assay (Figure 3.4B). Disulfide isomerase activity is thought to be associated with oligomerisation of thioredoxin-fold proteins (41,123,175) so we were interested to determine whether PmScsCΔLinker forms a trimer - like the full-length native protein - or a monomer - like the truncated PmScsCΔN variant.

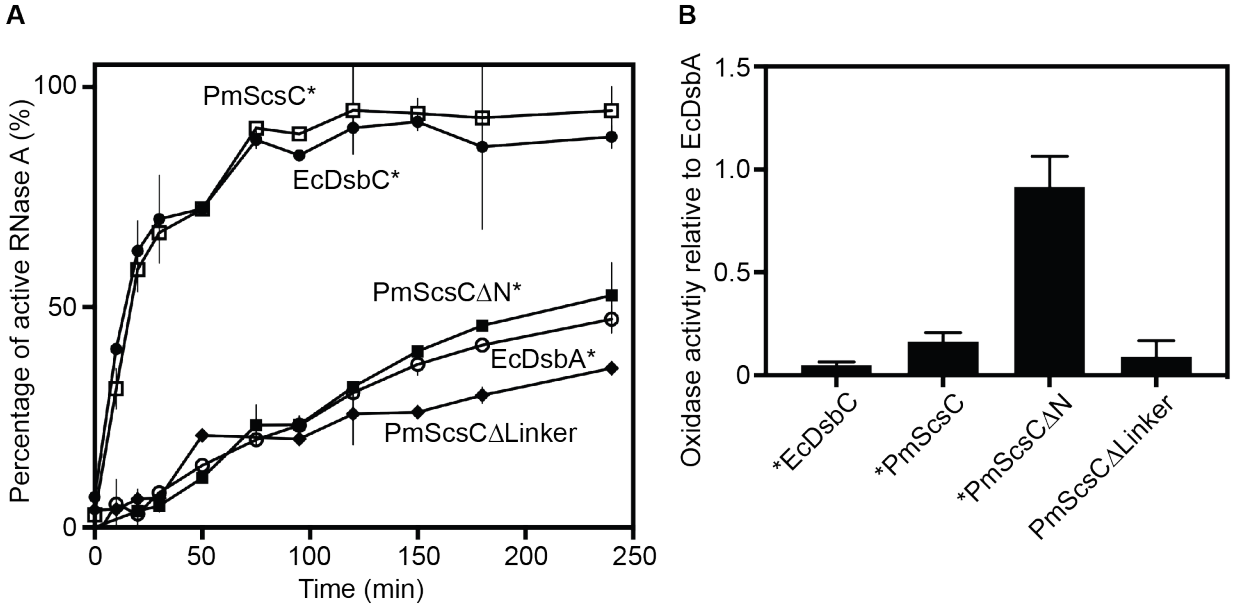


Figure 3.4: PmScsCΔLinker biochemical activity. A. *The ability of PmScsCΔLinker to isomerise the incorrect disulfide bonds in scrambled RNase A is poor – i.e. even lower than that of the dithiol oxidase EcDsbA.* ***B.*** *PmScsCΔLinker has poor dithiol oxidase activity; i.e. much less than monomeric PmScsCΔN and similar to that of protein disulfide isomerases EcDsbC and PmScsC. Asterisks indicate data published previously (175). Error bars on both panels represent the standard deviation of the mean.*

3.5.3 The crystal structure reveals a trimeric form of PmScsCΔLinker

The three crystal structures of native PmScsC reported previously show that the linker residues can adopt helical (extended conformation), strand (intermediate conformation) or loop (compact conformation) secondary structure (175). In the extended conformation crystal structure, the linker adopts a helical conformation thereby creating a very long helix (35 residues, approximately 9.5 turns, Figure 3.5A) extending from the trimerisation stalk, through into the catalytic domain of PmScsC.

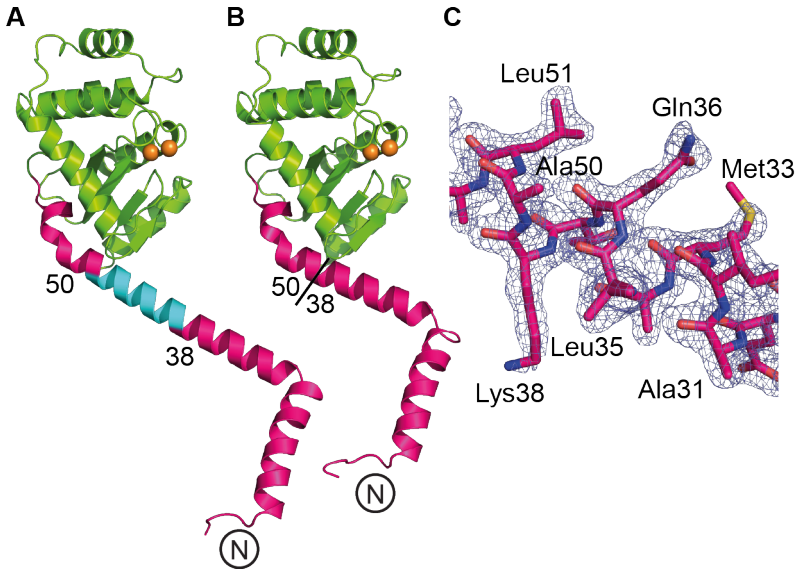


Figure 3.5: PmScsCΔLinker crystal structure. A. *Single protomer of the extended PmScsC (PDB: 5ID4).* ***B.*** *Single protomer of PmScsCΔLinker. The different features of the crystal structures are coloured as in Figure 3.1. The N-terminus of each structure is labelled and the positions of the residues flanking the flexible linker in the native structure are indicated.* ***C.*** *Electron density around the deleted region in the PmScsCΔLinker structure. The wire mesh represents the $2F_o$-F_c electron density map (generated in phenix.maps) contoured to 1σ and shown within a 2Å radius of each atom.*

We solved the crystal structure of PmScsCΔLinker at a resolution of 2.08 Å using molecular replacement (Table 3.1, Figure 3.5B, C). The variant crystallised in the same space group (H32) and crystallisation condition as the extended conformation PmScsC (PDB: 5ID4). However, the unit cell dimensions of PmScsCΔLinker are shorter than those of the extended PmScsC structure (64.0 Å, 64.0 Å, 299.8 Å compared to 86.7 Å, 86.7 Å, 330.0 Å). The crystal structure of PmScsCΔLinker – like the crystal structure of the extended form of PmScsC – is trimeric through crystallographic symmetry, with a single protomer in the asymmetric unit (Figure 3.6A). The regions on either side of the linker

region reported to be helical in the three native PmScsC crystal structures are directly connected in PmScsCΔLinker and are also helical.

The crystal structure of PmScsCΔLinker is similar to that of extended PmScsC except for a shortening of the long helix by almost three turns (Figure 3.5A, B). This shortened helix in the crystal trimer of PmScsCΔLinker brings the three catalytic domains into much closer proximity in the variant compared with the native extended conformation. This proximity is evident from the distance between the catalytic cysteines of protomers: 15 Å in PmScsCΔLinker and 34 Å in the extended native structure (measured between the Cβ atoms of C84, Figure 3.6A, B).

The PmScsCΔLinker crystal structure also reveals interactions between each of the catalytic domains of the trimer. The surface interface is formed between a positively charged patch close to the CXXC motif on the catalytic domain of one protomer and a neutral/hydrophobic patch on another (Figure 3.6C). This interface is not formed in the extended PmScsC structure, which has both of these regions exposed to solvent and it is also not formed by crystal packing in the PmScsCΔN structure.

The trimerisation stems of all three native PmScsC crystal structures are α-helical with a hydrophobic core, and hydrogen bonds between surface exposed polar/charged residues. These features are conserved in the crystal structure of PmScsCΔLinker (Figure 3.6D and E). However, it should be noted that PmScsCΔLinker has an additional unanticipated mutation (Asn6Lys) at the N-terminus. In the three native crystal structures the side chain of the Asn6 residue forms an intramolecular hydrogen bond to the main chain of Gln9. In the PmScsCΔLinker crystal structure, the side chain of Gln9 adopts the same conformation as in all three native PmScsC structures but the mutated residue Lys6 does not form an equivalent hydrogen bond with the Gln9 main chain. The lack of this specific hydrogen bond could potentially increase N-terminal disorder, though the crystal structure suggests there is little impact on the hydrogen bonds and hydrophobic contacts that contribute to trimerisation.

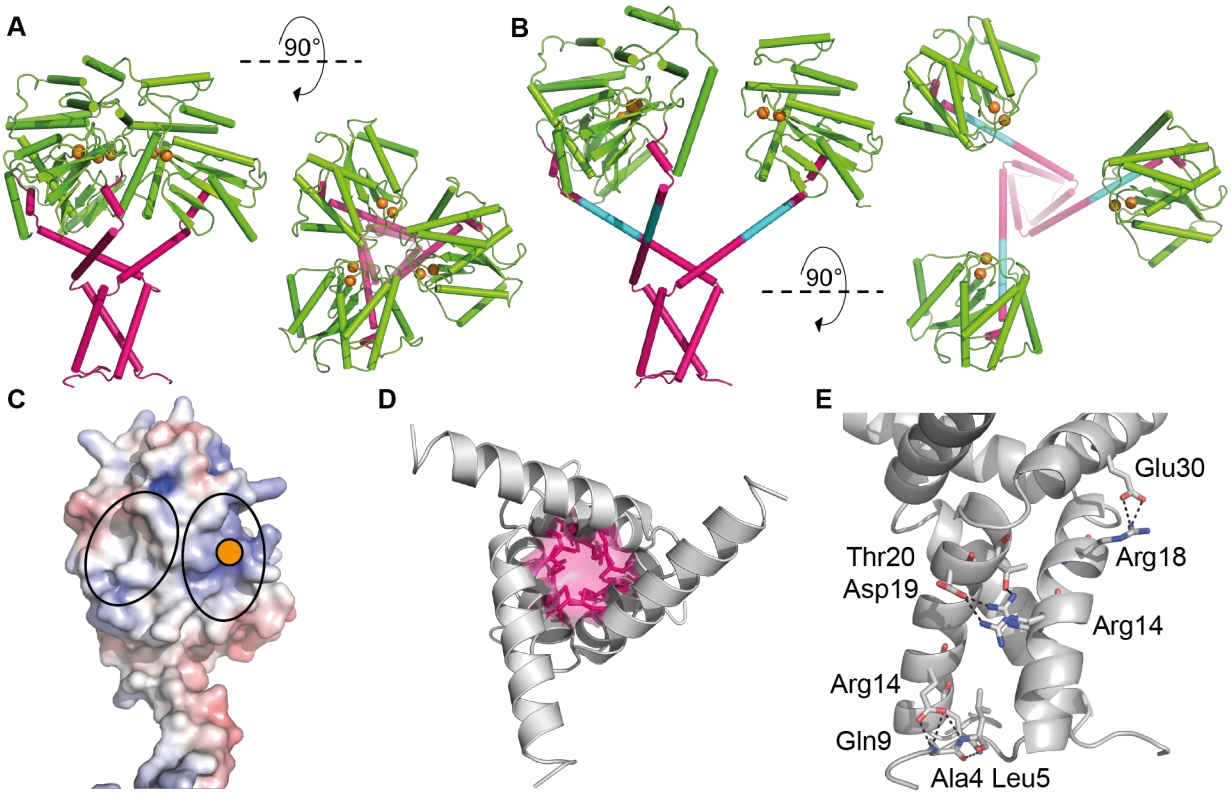


Figure 3.6: Trimerisation of PmScsCΔLinker. A. *Side and top views of the catalytic domain of the crystal structure of PmScsCΔLinker. The crystal structure reveals a trimer through crystallographic symmetry. In many respects this crystal structure resembles that of native extended PmScsC* ***B*** *(PDB: 5ID4). However, the extended helix linking the trimerisation domain to the catalytic domain is shorter in PmScsCΔLinker due to deletion of the linker (cyan in* ***B****). In both* ***A*** *and* ***B****, the proteins are coloured using the scheme from Figure 3.1.* ***C.*** *The electrostatic surface potential of the catalytic domain of the PmScsCΔLinker protomer structure. Electrostatic calculations were performed using APBS (11) and contoured to -7.5 (red) and +7.5 k_BT/e (blue). A solid orange circle indicates the position of the catalytic cysteines and black ovals surround the regions of the catalytic domains in close contact with adjacent catalytic domains in the PmScsCΔLinker crystal structure.* ***D.*** *The hydrophobic core of the trimerisation stem of PmScsCΔLinker; side chains of residues contributing to the hydrophobic core are coloured pink* ***E.*** *Amino acid residues mediating side chain polar and electrostatic interactions between the trimerisation stems of two PmScsCΔLinker protomers. Black dashed lines represent the hydrogen bonds and those residues involved are labelled.*

3.5.4 SAXS reveals that the oligomeric state of PmScsCΔLinker is concentration dependent

We had previously observed that the PmScsC RHP variant (in which the flexible linker is replaced with a rigid helix) forms a trimer - like the native protein - but is inactive in the protein disulfide isomerase assay (175). This lack of activity was thought to be a consequence of reduced catalytic domain motion observed in solution using small angle scattering (175). This finding led to the hypothesis that potent protein disulfide isomerase activity requires at least two catalytic domains (e.g. dimer, trimer) which have the ability to move independently of each other (175). Our present results show that PmScsCΔLinker also forms a trimer and is also inactive as a protein disulfide isomerase. We therefore anticipated that the lack of protein disulfide isomerase activity was a consequence of reduced catalytic domain rigid body motion, and sought to confirm this in solution using small angle scattering.

Unexpectedly, initial SAXS data for PmScsCΔLinker displayed a strong dependence on concentration, consistent with an equilibrium between different PmScsCΔLinker oligomeric forms (monomer, dimer, trimer). This result raised the possibility that, at low concentrations, the wild-type protein may also be present as a mixture of monomer, dimer and trimer, and this might also impact on its function or its regulation. To investigate this possibility further, SAXS data were collected from native PmScsC (Figure 3.7) and PmScsCΔLinker (Figure 3.8) at concentrations between 0.1 and 10.0 mg/mL (Table 3.2). At the lowest concentration (0.1 mg/mL), the shape of the $p(r)$ function for native PmScsC differed from the shapes of those at higher concentrations, indicating some dissociation of the wild-type protein. This effect was quantified by fitting a linear combination of monomer, dimer and trimer scattering to the data at each concentration (Figure 3.7 and Table 3.3). For PmScsCΔLinker, the results are very different; the shape of the $p(r)$ curve changes significantly with each concentration evaluated. Again, this effect was quantified by fitting a linear combination of monomer, dimer and trimer scattering to the data at each concentration (Figure 3.8 and Table 3.3). The relative amounts of each were constrained by common equilibrium constants. The analysis worked well, but was limited by the fact that the structures of monomer and dimer forms in solution are not known. In each case, we made a simplistic assumption that a monomer adopted the structure of a single protomer from the corresponding crystal structure, and a dimer could be represented by removing a single protomer from the trimer crystal structure. We also know that native PmScsC is dynamic in solution (175). While the fits we obtained of the model curves to the

scattering data are reasonable, there are systematic differences between the experimental data and the model, especially at higher concentrations. The limitations of the modeling (lack of accurate models, dynamic motion) explains most of this variation.

Overall, the SAXS analysis suggests that only 1% of native PmScsC is present as monomer in solution even at the lowest concentration of 0.1 mg/mL; we conclude that the trimer is the dominant form of the native protein in solution even at low concentrations. For the variant PmScsCΔLinker, it is a different story; the protein is in equilibrium between different oligomerisation states. At 0.1 mg/mL approximately 97% of the PmScsCΔLinker solution is monomer, 3% is dimer with negligible trimer. At 10 mg/ml these proportions change to 40% monomer, 48% dimer and 12% trimer.

Table 3.3: Oligomer fractions and equilibrium constants.

Measured Concentration (mg/mL)	Calculated Concentration (mg/mL)	Monomer Fraction	Dimer Fraction	Trimer Fraction
Native PmScsC ($K_t = 6.3 \times 10^{-18}$ M^2)				
0.1	0.05	0.01	-	0.99
0.5	0.33	0.00	-	1.00
2.0	1.48	0.00	-	1.00
5.0	3.74	0.00	-	1.00
10.0	7.07	0.00	-	1.00
PmScsCΔLinker ($K_d = 2.1 \times 10^{-4}$ M; $K_t = 7.30 \times 10^{-4}$ M)				
0.1	0.07	0.97	0.03	0.00
0.5	0.36	0.88	0.12	0.00
2.0	1.31	0.71	0.27	0.02
5.0	3.55	0.53	0.41	0.07
10.0	7.17	0.40	0.48	0.12

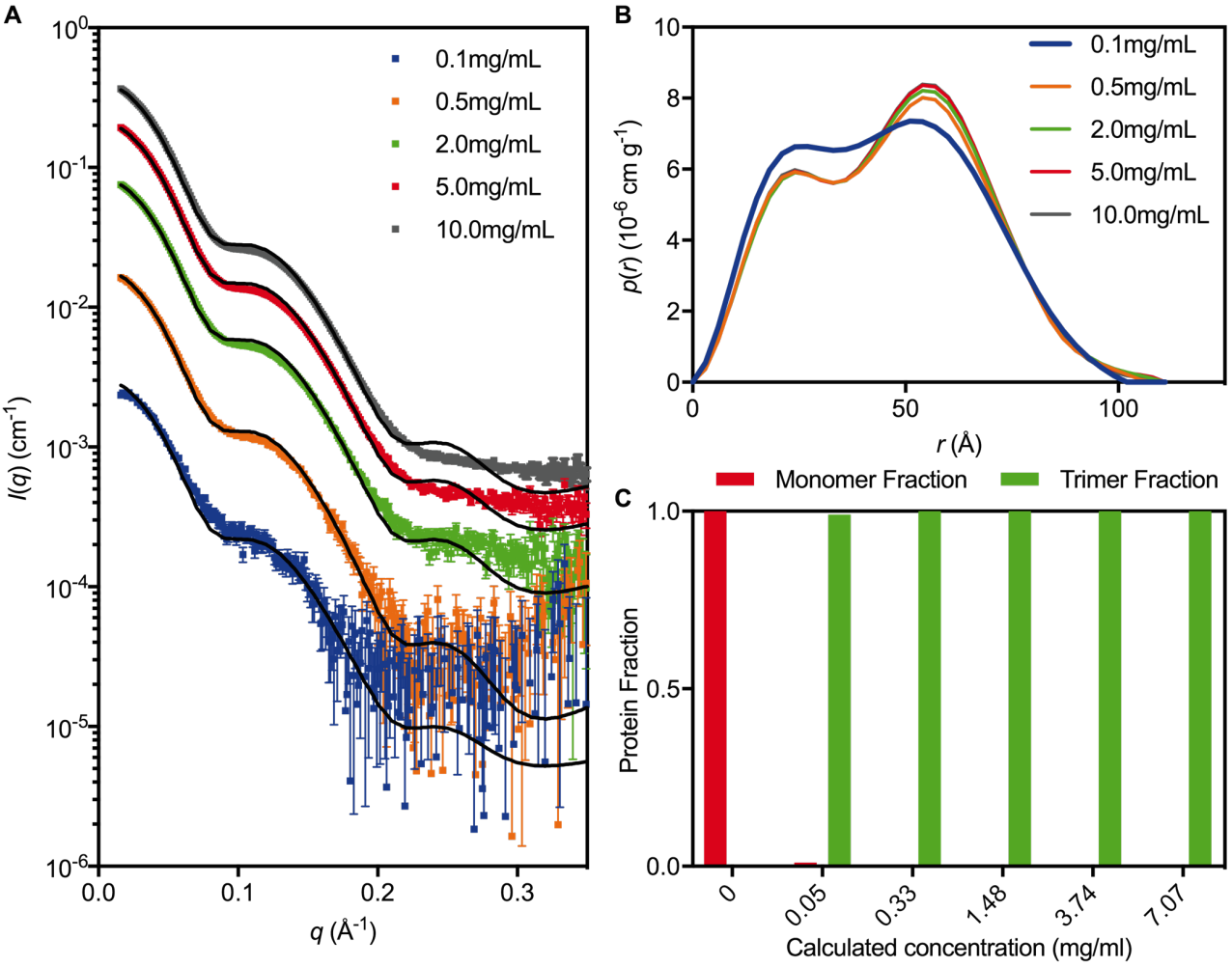


Figure 3.7: Native PmScsC SAXS data. A. *Scattering data collected at ~0.1 mg/mL (blue, χ^2 = 2.55), ~0.5 mg/mL (orange, χ^2 = 7.05), ~2.0 mg/mL (green, χ^2 = 72.85), ~5.0 mg/mL (red, χ^2 = 371.94), ~10.0 mg/mL (grey, χ^2 = 815.23). The overlaid line (black) on each scattering curve is the fitted linear combination of monomer and trimer scattering curves.* ***B.*** *The pair-distance distribution function, p(r), derived from the scattering data normalised by concentration show the 0.1 mg/mL data set differs slightly from that of the other concentrations (which all overlay), indicating the possibility that monomer is present in solution at the lowest concentration.* ***C.*** *The fraction of protein present as monomer and trimer, derived from the fitted models, shows that at total protein concentrations above 10 μM, the protein is almost exclusively present as a trimer in solution.*

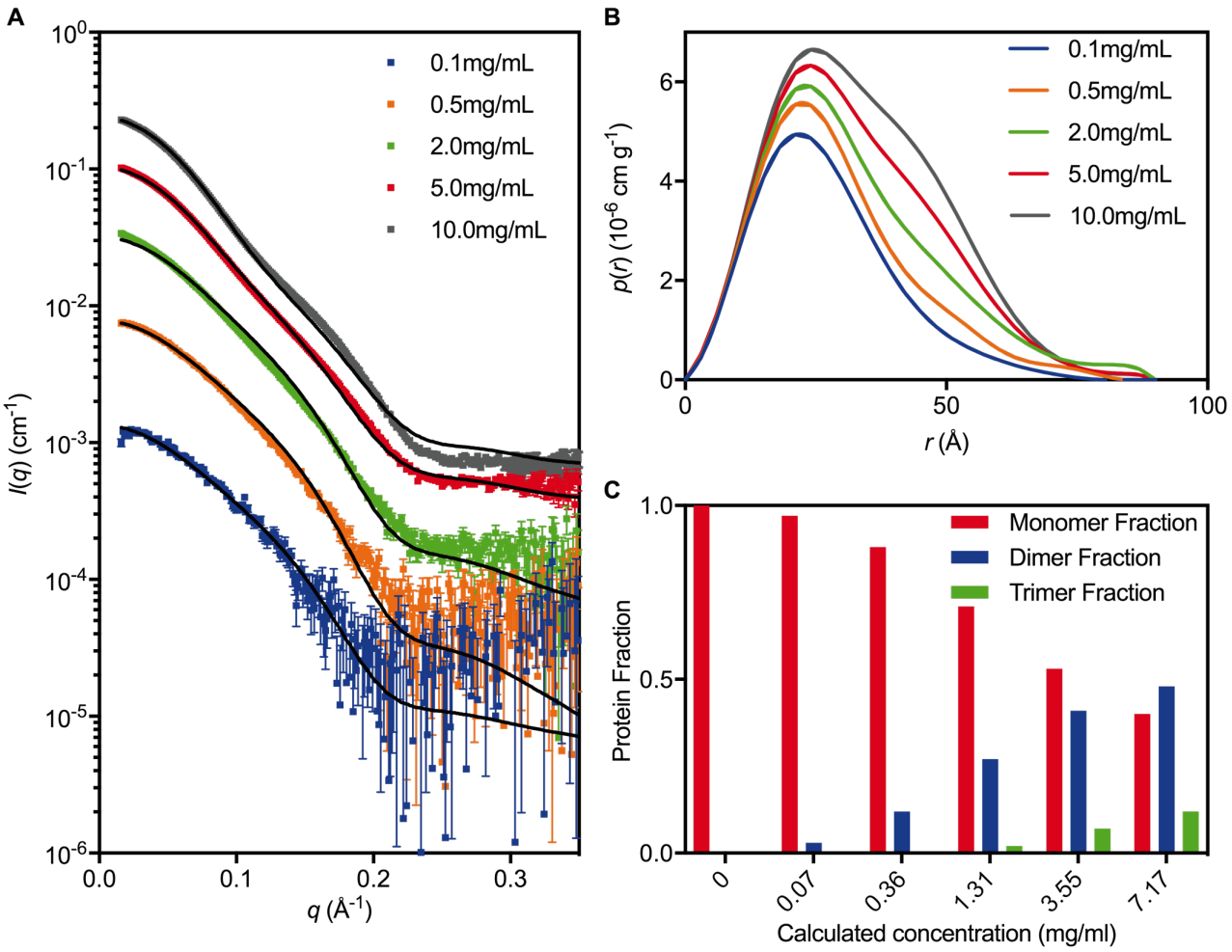


Figure 3.8: PmScsCΔLinker SAXS data. A. *Scattering data collected at ~0.1 mg/mL (blue, χ^2 = 1.27), ~0.5 mg/mL (orange, χ^2 = 2.19), ~2.0 mg/mL (green, χ^2 = 6.89), ~5.0 mg/mL (red, χ^2 = 37.59), ~10.0 mg/mL (grey, χ^2 = 177.29). The overlaid line (black) on each scattering curve is the fitted linear combination of monomer and trimer scattering curves.* ***B.*** *The pair-distance distribution function, p(r), derived from the scattering data normalised by concentration show that the oligomeric composition of the solution is highly dependent on concentration.* ***C.*** *The fraction of protein present as monomer, dimer, and trimer, derived from the fitted models, shows the monomeric species is dominant in solution until the total protein concentration reaches ~240 μM. The concentration of trimer rises slowly, and reaches a protein fraction of ~0.2 at 1.7 mM (40 mg/mL), which is the concentration at which the crystallisation was performed.*

3.6 Discussion

This study focused on characterising two variants of PmScsC, which - in its native form - is a highly dynamic, trimeric disulfide isomerase. The native PmScsC protomer consists of an N-terminal trimerisation stem with an 11 amino acid flexible linker connected to a C-terminal thioredoxin fold catalytic domain harbouring a CXXC active site. To better understand the role of the trimerisation stem and flexible linker we designed the variants PmScsCΔN and PmScsCΔLinker. PmScsCΔN lacks the first 41 amino acids of the native PmScsC construct; its loss of disulfide isomerase activity and gain of dithiol oxidase activity has been reported previously (175). PmScsCΔLinker is newly reported here, and lacks the 11 amino acids that form the flexible linker. We reported here the structure of PmScsCΔN and analysed the structure and biochemical function of PmScsCΔLinker.

The X-ray crystal structure of PmScsCΔN agrees with previous MALLS data showing that this variant is monomeric (175). The crystal structure is also consistent with solution SAXS data and the low-resolution model obtained from SAXS analysis. Although the PmScsCΔN crystal structure is very similar to the structures of the catalytic domains in each of the three native PmScsC crystal structures, unlike the native enzyme it lacks disulfide isomerase (disulfide shuffling) activity but instead has dithiol oxidase (disulfide forming) activity. The PmScsC catalytic domain has an acidic catalytic cysteine and an energetically unfavourable disulfide (175) – that are also features of the archetypal but monomeric dithiol oxidase EcDsbA – so it is perhaps not surprising that removal of the N-terminal trimerisation residues leads to oxidase activity in a model assay. Moreover, the structural similarity between monomeric PmScsCΔN and the catalytic domains of native trimeric PmScsC supports the notion that it is the N-terminal stem – defining whether the catalytic domain is monomeric or trimeric - that is the determinant of isomerase or oxidase activity. Put simply, the differential activity we observe between native PmScsC and variant PmScsCΔN is not a specific feature of the catalytic domain because the same catalytic domain has the propensity for oxidase (monomeric protein) or isomerase (trimeric protein) activity depending on its context. The isomerase activity of PmScsC requires both its N-terminal trimerisation domain and its C-terminal catalytic domain, in the same way that the isomerase activity of EcDsbC requires its dimerisation domain and its catalytic domain (41).

Our SAXS analysis showed that native PmScsC is a trimer in solution, even at low concentrations. However, SAXS analysis of PmScsCΔLinker showed that it exists

predominately as a monomer or a low affinity dimer (K_d ~ 200 μM) in preference to a trimer in solution even though it crystallises in the trimeric form. Together these data suggest that removal of the flexible linker of PmScsC between the trimerisation stem and catalytic domains has the unintended consequence of reducing the propensity to form a stable trimer in solution.

The crystal structure of PmScsCΔLinker provides insight into why the trimeric form of the variant may be energetically unfavourable compared with the monomeric and dimeric forms in solution. The trimerisation stem interactions of PmScsCΔLinker are similar to those observed in the native enzyme crystal structures, but all three catalytic motifs are in closer proximity to each other than in any of the native PmScsC crystal structures. This close packing suggests the possibility of unfavorable contacts between catalytic domains in PmScsCΔLinker that may contribute to the instability of the trimer. Indeed, the PmScsCΔLinker crystal structure reveals a close association of a positively charged patch near the catalytic motif on the accessible surface of one protomer with a neutral/hydrophobic patch on another. In the crystal packing of the PmScsCΔN structure (which does not have the constraint of the trimerisation stem), there are no contacts involving the positively charged surface patch near the catalytic motif, suggesting that packing against this site is not favoured. Thus, the formation of the PmScsCΔLinker trimer is likely to be dependent on the balance of favourable interactions formed between the trimerisation stems and unfavourable interactions between the catalytic domains. At lower concentrations the unfavourable interactions between the catalytic domains may override the favourable interactions but at higher concentrations, with more molecules in close contact, it may be more thermodynamically favourable for the hydrophobic portion of the trimerisation helices to interact with each other. This interplay between interactions of the catalytic domains and trimerisation stems could explain why the oligomeric state of PmScsCΔLinker in solution alters with concentration of the protein. Taken together, the crystal structure and SAXS analysis of PmScsCΔLinker suggests that the 11 amino acid linker plays a dual role by (i) acting as a spacer between the trimerisation and catalytic domains to enable formation of a stable trimer and (ii) to provide the necessary functional dynamics and cooperativity between the catalytic domains (175).

Removal of the 11 amino acid flexible linker from PmScsC impacts on the activity of the protein: unlike the native protein, the PmScsCΔLinker has no disulfide isomerase activity. Structural characterisation of PmScsCΔLinker revealed two potential reasons for this

result. Firstly, at the concentration used in the assay (10 μM) it is likely that most of the protein would be monomeric (Figure 3.8C), and we know from PmScsCΔN variant studies that monomeric protein has low protein disulfide isomerase activity (175). Secondly, even if there was sufficient trimeric PmScsCΔLinker present in solution, the crystal structure suggests there would be limited space between the catalytic domains to bind a misfolded protein substrate, and possibly limited motion to allow disulfides to be shuffled. It is interesting to note that PmScsCΔLinker is inactive in the dithiol oxidase assay. Given that the protein would be predominantly monomeric at the concentration tested in this assay (80 nM, Figure 3.8C) we expected that, like monomeric PmScsCΔN, it would have oxidase activity. One explanation for this result is that the N-terminal hydrophobic regions of PmScsCΔLinker somehow occlude the substrate-binding site that is available in monomeric PmScsCΔN.

Overall, our structural and functional characterisation of PmScsCΔN and PmScsCΔLinker provides new insight into the importance of unique structural features of the trimeric disulfide isomerase PmScsC. We confirmed previous assumptions that disulfide isomerases are similar to but differentiated from dithiol oxidases by requiring both an oligomerisation domain and a catalytic domain. We further propose that the flexible linker of PmScsC is important not just as a shape-shifting peptide, but also as a spacer, to enable functional flexibility between the catalytic domains. We suggest that this functional flexibility between the catalytic domains is necessary for cooperative enzymatic activity of disulfide isomerase PmScsC. Finally, on a more general note, we found that deleting amino acids from a protein can have unintended consequences. In this case, deletion of the flexible linker unexpectedly changed the oligomerisation kinetics of the protein, even though the region directly involved in trimerisation remained intact.

3.7 Acknowledgments

We acknowledge the use of the University of Queensland Remote Operation Crystallisation and X-ray (UQ ROCX) facility and the support from staff, Gordon King and Karl Byriel. We also acknowledge use of the Australian Synchrotron MX and SAXS facilities and thank the staff for their support. This work was supported by an Australian Research Council Laureate Fellowship (FL0992138, to J.L.M.), an Australian Government Research Training Program stipend (to E.J.F.) and an Institute for Molecular Bioscience Research Advancement Award (to E.J.F.).

Chapter 4

4. *P. mirabilis* ScsBα and ScsC form a functional redox relay

4.1 Chapter Foreword

This chapter is a reproduction of a research article that was originally published in the *Journal of Biological Chemistry*. It contains the crystal structure of the N-terminal domain of PmScsB (PmScsBα), confirms that this domain is the redox partner of the trimeric disulfide isomerase PmScsC and shows a SANS derived model of the PmScsC-PmScsBα interaction. Minor changes have been made to the paper to conform to the thesis submission requirements. These changes include, the overall formatting, removal of the keywords and contributions (which can be found in the preliminary pages) and the numbering of tables, figures and references.

4.1.1 Reference

Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E. and Martin, J. L. (2018) Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB. *Journal of Biological Chemistry* 293(16): 5793-5805; doi: 10.1074/jbc.RA118.001860. © The Authors.

4.1.2 Contribution

I made a significant contribution to the experimental design of the work outlined in this publication, the analysis of results and the production of the publication. I designed the cysteine variants and chose the assays to use to investigate the interaction between PmScsC and PmScsBα. I also optimised the PmScsBα crystallisation, collected and processed diffraction data and solved and refined the structure of this protein. I performed the interaction assays between PmScsC and PmScsBα and analysed the data. I created the cysteine mutants of PmScsC and PmScsBα and produced the 3:1 PmScsC-PmScsBα complex for SANS analysis. I wrote the first draft of the paper and designed and created the figures.

Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB

Emily J. Furlong[1,2], Hassanul G. Choudhury[1#], Fabian Kurth[1&], Anthony P. Duff[3], Andrew E. Whitten[3*] and Jennifer L. Martin[1,2*]

From the [1]Institute for Molecular Bioscience, University of Queensland, St Lucia, QLD 4072, Australia. [2]Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD 4111 Australia. [3]Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW 2234, Australia.

Running title: *P. mirabilis ScsBα and ScsC form a functional redox relay*

Present address: [#]Cello Health Consulting, Farnham Surrey GU9 7DN, UK. [&]Bristol-Myers Squibb, Arnulfstraße 29, 80636 Munich, Germany.

*To whom correspondence should be addressed: Jennifer L. Martin: Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD 4111, Australia; jlm@griffith.edu.au or Andrew E. Whitten: Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW 2234, Australia; awh@ansto.gov.au.

Publication timeline

Received for publication: 12 January 2018

Received revised form: 15 February 2018

Published (Papers in Press): 28 February 2018

Final version published: 20 April 2018

4.2 Abstract

Correct disulfide bond formation is essential for proper folding of many proteins, including bacterial virulence factors. The suppressor of copper sensitivity (Scs) proteins have roles in dithiol/disulfide interchange and the bacterial response to copper stress. Encoded in a four-gene cassette (ScsABCD) present in many Gram-negative bacteria, the Scs proteins are enigmatic and poorly characterised. Here, we show that the periplasmic α domain of the membrane protein ScsB in the Gram-negative bacterium *Proteus mirabilis* forms a redox relay with the soluble periplasmic protein PmScsC. We also found that the periplasmic α domain is sufficient to activate the disulfide isomerase activity of PmScsC. The crystal structure of PmScsBα at a resolution of 1.54 Å revealed that it comprises two structurally similar immunoglobulin-like folds, one of which includes a putative redox-active site with the sequence CXXXC. We confirmed the importance of these cysteine residues for PmScsBα function and in addition, we engineered cysteine variants that produced a stable complex between PmScsC and PmScsBα. Using small-angle X-ray and neutron scattering analyses with contrast variation, we determined a low-resolution structure of the PmScsC-PmScsBα complex. The structural model of this complex suggested that PmScsBα uses both of its immunoglobulin-like folds to interact with PmScsC and also revealed that the highly dynamic PmScsC becomes ordered upon PmScsBα binding. These findings add to our understanding of the poorly characterised Scs proteins.

4.3 Introduction

The correct formation of disulfide bonds is an essential component in the folding of many proteins, including bacterial virulence factors. In bacteria, disulfide bond forming (Dsb) proteins are responsible for introducing disulfide bonds into substrate proteins (DsbA and B) (7,25) as well as reducing and isomerising (proof-reading and shuffling) disulfide bonds that have been incorrectly introduced (DsbC and D) (39,44). Dsb proteins function through the redox action of two catalytic cysteines that are often embedded in a thioredoxin fold. These systems are well characterised in *E. coli*; however, homologues of Dsb proteins exist in a range of Gram-negative bacteria. A group of related and poorly studied proteins found most commonly in Proteobacteria are the suppressor of copper sensitivity (Scs) proteins (118). The Scs proteins contribute to the bacterial virulence trait of resistance to copper stress (111,120). Originally identified in *Salmonella enterica* serovar Typhimurium (111), the *scs* locus encodes four proteins ScsA-D, all of which have predicted catalytic motifs consisting of two cysteines (CXXC or CXXXC). Three of these (ScsB-D) incorporate a predicted thioredoxin fold and one of the four (ScsA) has a predicted copper-binding

motif. ScsD has a predicted N-terminal membrane anchor linked to a predicted periplasmic thioredoxin-fold domain (111), reminiscent of proteins involved in cytochrome c biogenesis such as DsbE/CcmG (69,124). The best studied of the four Scs proteins is the soluble periplasmic protein, ScsC (118,120,175). An extraordinary feature is the variation in structure and function of ScsC proteins across different bacteria. *Caulobacter crescentus* ScsC (CcScsC) is reported to be a dimeric disulfide isomerase (118), *Salmonella enterica* serovar Typhimurium ScsC (StScsC) is a monomeric redox protein with no disulfide isomerase activity (120) and *Proteus mirabilis* ScsC (PmScsC) is a highly dynamic "shape-shifting" trimeric disulfide isomerase (175).

The present work focuses on the structure and function of *P. mirabilis* ScsB, a putative redox partner of *P. mirabilis* ScsC. ScsB resembles *E. coli* DsbD in many respects (118). EcDsbD is the membrane protein partner of the archetypal disulfide isomerase *E. coli* DsbC. Like EcDsbD, the ScsB protein is predicted to comprise three domains: an N-terminal periplasmic domain (DsbDα/ScsBα), a transmembrane domain (DsbDβ/ScsBβ) of eight helices; and a periplasmic C-terminal thioredoxin-fold domain (DsbDγ/ScsBγ) (Figures 4.1A, B). Each of the three domains in both EcDsbD and ScsB have two catalytic cysteine residues. In DsbD, these cysteine pairs in each domain form a reduction cascade that originates from thioredoxin reductase (TR) and NADPH in the cytoplasm, is passed through thioredoxin to DsbDβ, DsbDγ, and then DsbDα, which reduces the cysteines in specific substrate proteins, such as DsbC (60,191) (Figure 4.1A). The similarity between the predicted architectures of EcDsbD and ScsB suggests the ScsB protein forms a similar redox relay system. Indeed, *C. crescentus* ScsB has been shown to interact with CcScsC and to maintain CcScsC in its reduced state, which is necessary for disulfide isomerase activity (118). A major difference between EcDsbD and ScsB is the size of their N-terminal α-domains. DsbDα comprises ~165 residues whereas ScsBα is much larger at ~255 residues, and there is no detectable sequence identity between the α-domains of DsbD and ScsB. Since the α-domain of EcDsbD interacts directly with substrates - such as EcDsbC - the different size and sequence of the ScsB α-domain suggests that the proteins interact with different substrates (118).

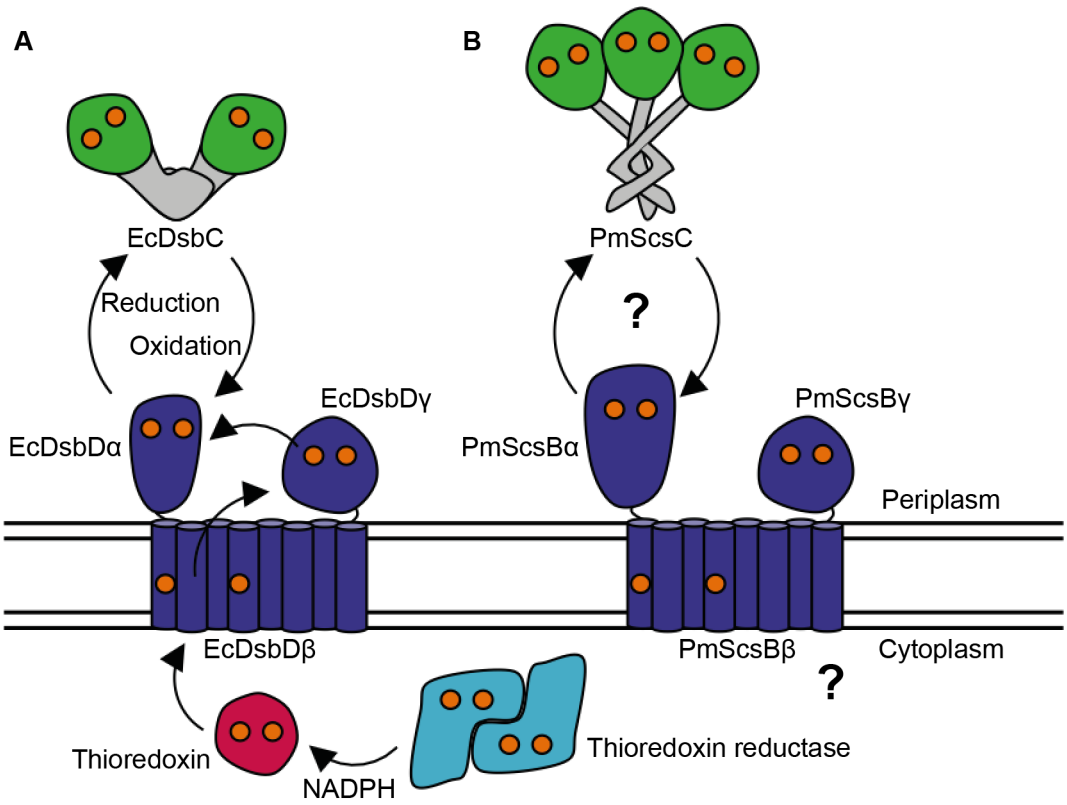


Figure 4.1: Schematic representation of bacterial disulfide isomerases and their redox relay partners. A. *EcDsbC is reduced by EcDsbDα, which obtains its reducing power from a reduction cascade that starts with thioredoxin reductase in the bacterial cytoplasm.* ***B****. PmScsC and PmScsBα are putative interaction partners and the source of PmScsB's reducing power is yet to be determined (represented by the question marks). The catalytic cysteines in each protein are shown as orange spheres.*

In the present study we investigated the structure and function of *P. mirabilis* ScsBα, showing that it forms a functional redox relay with the highly dynamic, trimeric disulfide isomerase PmScsC. We also report the first crystal structure of any ScsBα, revealing unexpectedly the presence of tandem immunoglobulin folds in PmScsBα, and showing that this same arrangement is likely to be shared in ScsBs encoded by other organisms. We created a stable PmScsC-PmScsBα (trimer-monomer) complex, and analysed small-angle X-ray scattering (SAXS) data and small-angle neutron (SANS) data with contrast variation to produce a low-resolution model of the complex. This model revealed that the highly dynamic PmScsC protein becomes more symmetric and ordered upon interaction with PmScsBα and that both immunoglobulin folds of PmScsBα interact with one protomer of PmScsC.

4.4 Results

4.4.1 PmScsC and PmScsBα form a specific redox relay system

From sequence analogy with the *C. crescentus* proteins CcScsB and CcScsC, PmScsBα is predicted to be the redox partner of PmScsC. If this functional relationship holds true, PmScsBα would specifically reduce oxidised PmScsC to convert it to the active reduced form. To confirm that a redox relay occurs between PmScsC and PmScsBα, we performed a redox gel-shift assay. First, we mixed PmScsC with PmScsBα and then analysed the redox states of the two proteins after an incubation period. The redox state was defined by analysing the proteins on SDS-PAGE after reaction with 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS) (Figure 4.2A). When AMS binds to a reduced thiol, the mass is increased by 0.5 kDa thus allowing discrimination between reduced and oxidised cysteine forms of a protein on mass-based gel separation. The results showed that when oxidised PmScsC was incubated with reduced PmScsBα, both proteins changed their redox states, confirming that disulfide exchange occurs between the two proteins (Figure 4.2B). The reverse redox reaction did not occur: reduced PmScsC remained reduced when mixed with oxidised PmScsBα. Moreover, we were able to confirm the residues involved in the change in redox state because addition of the variant PmScsBα C114A to oxidised PmScsC did not catalyse disulfide exchange. However, there was some evidence for formation of a PmScsC-PmScsBα C114A complex (band marked X, Figure 4.2B).

Control experiments using PmScsC or PmScsBα alone showed that, under the conditions of the experiment, these proteins did not spontaneously change their redox state. These results support the notion that PmScsC and PmScsBα form a specific redox relay system (Figure 4.2B) and that the residues C114 and C118 of PmScsBα are responsible for the redox interaction.

4.4.2 PmScsBα primes the disulfide isomerase activity of PmScsC

We have previously shown that PmScsC is a disulfide isomerase (175). We tested the hypothesis that the redox interaction between PmScsBα and PmScsC might contribute to this function of PmScsC. We measured the disulfide isomerase activity of PmScsC using scrambled RNase A as substrate (Figure 4.2A). Reduced PmScsC, but not oxidised PmScsC, was able to refold scrambled RNase A, and thereby activate the enzyme (175) (Figure 4.2C). PmScsBα on its own (oxidised or reduced) had no effect on the activity of scrambled RNase A. However, in the presence of reduced PmScsBα the inactive oxidised

PmScsC was able to restore the activity of RNase A. This result supports the hypothesis that PmScsBα reduces the disulfides of oxidised PmScsC, thereby enabling its disulfide isomerase activity (Figure 4.2D). This experiment also confirmed that disulfide exchange in this system does not flow in the reverse direction: reduced PmScsC did not lose its ability to shuffle scrambled RNase A disulfides when it is incubated with oxidised PmScsBα. This may suggest that reduced PmScsC is unable to interact with PmScsBα. Further, the cysteine variant in which C114 is mutated to Ala (PmScsBα C114A) does not activate oxidised PmScsC, showing that this cysteine is required for the redox relay to operate.

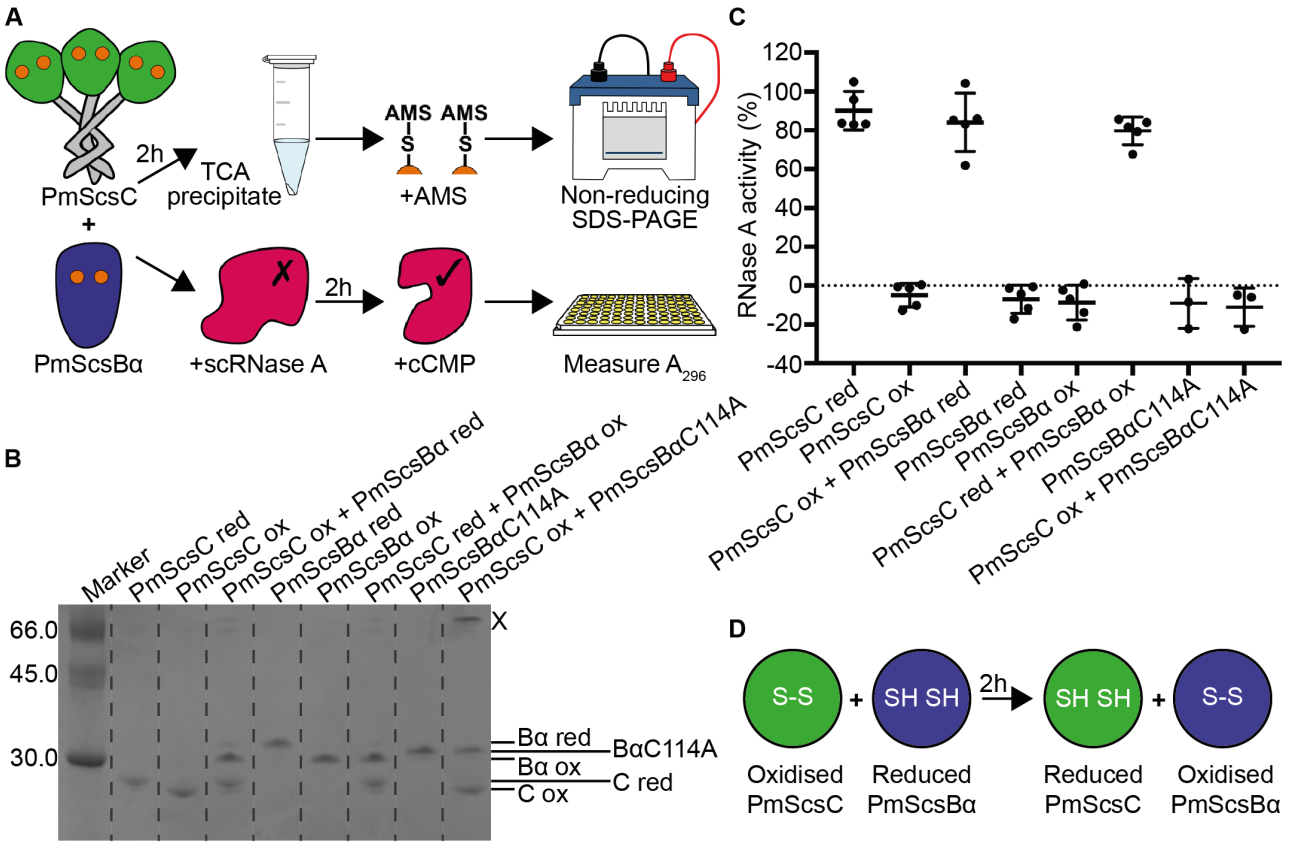


Figure 4.2: PmScsBα reduces and activates PmScsC. A. *Schematic diagrams of the gel-shift and scRNase A experiments.* ***B.*** *A representative gel from the gel shift assay showing the interaction between PmScsC and PmScsBα. Dotted lines are a visual aide to delineate lanes on the gel. The band marked X is likely to be 1:1 PmScsC-PmScsBα C114A complex, as the usually trimeric PmScsC runs as a monomer on SDS-PAGE.* ***C.*** *Results from the modified scrambled RNase A assay. Activity is shown relative to native RNase A (100%) and scrambled RNase A only (0%) controls. For each sample, the activity of three or five replicates and the mean with s.d. is shown.* ***D.*** *Schematic showing the redox interaction that occurs between PmScsC and PmScsBα.*

4.4.3 Crystal structure of PmScsBα

After removal of the periplasmic signal sequence, the predicted PmScsBα domain comprises 255 residues (residues 21-275, of the reported PmScsB protein sequence B4EV20). The crystal structure of this domain was solved to a resolution of 1.54 Å by using SeMet-labelled protein together with SAD phasing (PDB: 6C29, Table 4.1, Figures 4.3A-D). The structure revealed a monomer consisting of two immunoglobulin-like folds linked by an α-helix. This arrangement was consistent across all three molecules in the asymmetric unit (RMSD 0.7-1.0 Å for 247 aligned residues). Curiously, the two subdomains A and B are reasonably similar in structure (Figure 4.3D RMSD 2.5 Å for 81 aligned residues), though the sequence identity for such an alignment is very low (~7%). The subdomains are each comprised of two β-sheets, one with three β-strands and the other with four β-strands (Figure 4.3A and B). Subdomain A has additional features, including: the cysteine motif ($C_{114}XXXC_{118}$); two β-strands forming a β-hairpin that links the two β-sheets of the canonical immunoglobulin fold; and a loop between β2 and β3, located on one side of the catalytic cysteines (arrows, Figure 4.3D). The cysteines are present in the reduced form in all three molecules in the asymmetric unit.

The presence of an immunoglobulin-like fold in both PmScsBα subdomains resulted in many hits in a DALI search against the Protein Data Bank (192). Notably, the top hit was the functionally equivalent *E. coli* DsbDα (PDB: 1JPE, RMSD 2.5 Å, 103 residues aligned), though this has only one immunoglobulin-like fold, which can be aligned with subdomain A of PmScsBα (Figure 4.4A). Like PmScsBα subdomain A, *E. coli* DsbDα has two catalytic cysteines and a β-hairpin close to the active site, but it does not have the β2-β3 loop of PmScsBα. Two other hits in the DALI search were the mouse γ-adaptin appendage domain from clathrin-binding adaptor AP-1 (PDB: 2A7B, RMSD 2.4 Å, 106 residues aligned) and the GAE domain of the clathrin-binding adaptor GGA protein from *Saccharomyces cerevisiae* (PDB: 3MNM, RMSD 2.4 Å, 102 residues aligned) (Figure 4.4B and C). These proteins have typical immunoglobulin-like folds, with no catalytic cysteines, and like EcDsbDα have one immunoglobulin fold that aligns with subdomain A of PmScsBα. These three proteins were also hits in a DALI search against PmScsBα subdomain B though with much higher RMSDs (1JPE: RMSD 3.8 Å, 75 residues aligned. 2A7B: RMSD 3.0 Å, 84 residues aligned. 3MNM: RMSD 2.9 Å, 86 residues aligned). Curiously, there were no high-scoring DALI hits that had a tandem immunoglobulin-fold arrangement like that of PmScsBα. We did find lower scoring DALI hits with two immunoglobulin-like folds (mostly antigen-binding fragments of antibodies e.g. PDB: 5JUE,

Z-score: 4.6, RMSD: 3.0 Å, 80 residues aligned), but they lacked the connecting α-helix of PmScsBα.

Table 4.1: Crystallography statistics for PmScsBα.

Data collection	**PDB: 6C29 (SeMet)**
Wavelength (Å)	0.9786
Resolution (Å)	97.06 – 1.538 (1.544-1.538)[a]
Space group	$P2_12_12_1$
Unit cell dimensions	
a, *b*, *c* (Å)	69.98, 97.06, 110.20
α, β, γ (º)	90, 90, 90
Number measured reflections	1761650
Number of unique reflections	111511
R_{merge}	0.146 (2.114)
$R_{p.i.m.}$	0.038 (0.531)
Mean <I/σI>	16.0 (2.1)
Redundancy	15.8 (16.3)
Completeness (%)	99.7 (98.3)
Wilson-B	15.96
SAD Phasing	
No. of sites	14
Figure of Merit	0.42
Refinement	
Number of monomers in a/u	3
Resolution used in refinement (Å)	55.10 – 1.538
Number of reflections	111388
R_{free} (%)/R_{work} (%)	19.8/17.3
Number of protein atoms	5860
Number of ligand atoms	0
Number of waters	669
B factors (Å^2)	
Average	25.5
Protein atoms	24.8
Ligands	NA
Waters	31.8
RMSD Bond length (Å)/ angles (º)	0.011/1.332
Ramachandran favoured / outlier (%)	97/0

[a]Values in parentheses refer to the highest resolution shell.

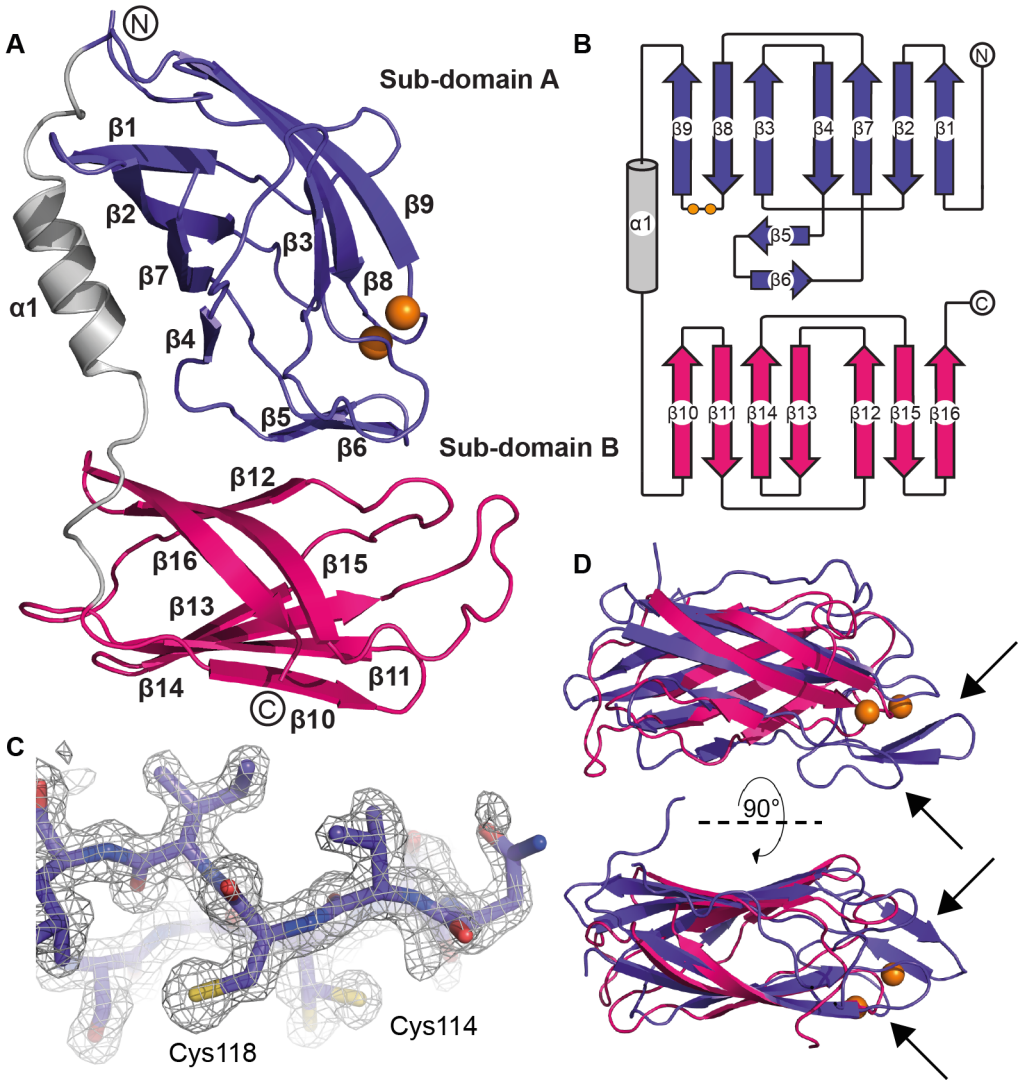


Figure 4.3: Structural characterisation of PmScsBα. A. *Cartoon representation of the crystal structure of PmScsBα (PDB: 6C29).* ***B.*** *Topology diagram of PmScsBα, showing the secondary structure features of the protein.* ***C.*** *The OMIT electron density around the catalytic loop in chain A of the PmScsBα crystal structure is displayed. The wire mesh represents the likelihood-weighted F_o – F_c electron density difference map, which is contoured to 3σ and was generated using phenix.refine with the region of interest omitted from the structure.* ***D.*** *Superimposition of subdomains A and B structures showing the structural similarity between the domains. Arrows indicate the inserted β-hairpin structure and the extended loop of subdomain A. In all panels, subdomain A is coloured purple and subdomain B is magenta, with the linking α-helix shown in grey. The catalytic cysteines are represented as orange spheres. The N- and C-termini are labelled in* **A** *and* **B**.

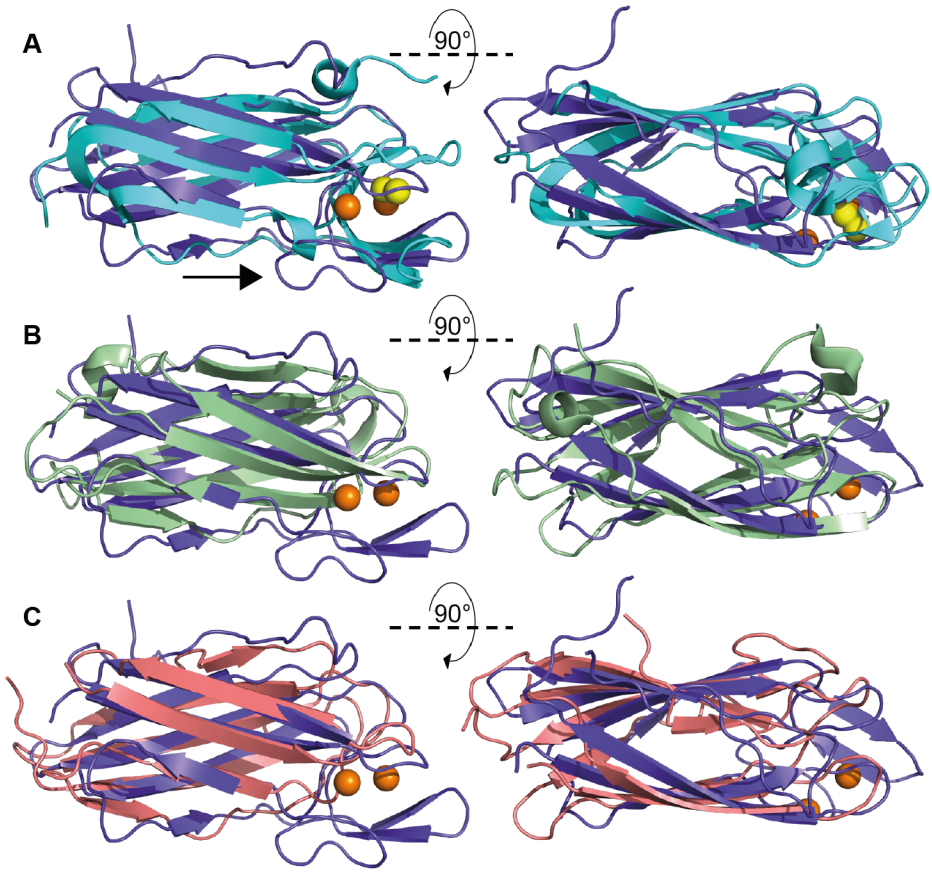


Figure 4.4: Structural overlay of PmScsBα subdomain A with selected DALI hits. A. *PmScsBα subdomain A (purple) aligned with EcDsbDα (PDB: 1JPE) (cyan). The catalytic cysteines of both PmScsBα subdomain A and EcDsbDα are shown as spheres and coloured in orange and yellow, respectively. An arrow indicates the extended loop in the PmScsBα structure.* ***B.*** *PmScsBα subdomain A (purple) aligned with PDB: 2A7B (light green).* ***C****. PmScsBα subdomain A (purple) aligned with PDB: 3MNM (light pink). Neither 2A7B nor 3MNM have catalytic cysteines so in both B and C the orange spheres correlate to the cysteines of PmScsBα.*

4.4.4 Structure-based sequence alignment

The amino acid sequence of PmScsBα was aligned with that of EcDsbDα using structure-based alignment, and the sequences of *C. crescentus* ScsBα (CcScsBα) and *S.* Typhimurium ScsBα (StScsBα) (Figure 4.5). ScsBα from *S.* Typhimurium and *C. crescentus* were chosen, as these are the only other organisms where the interaction partner, ScsC, has been characterised. EcDsbDα was chosen because of the known structural and functional relationship. PmScsBα shares 44% sequence identity with StScsBα (280 residues), 21% identity with CcScsBα (291 residues) and just 12% identity with EcDsbDα (116 residues of PDB: 1JPE). There are several conserved residues across the three ScsBα and EcDsbDα protein sequences aside from the catalytic cysteines. Other

conserved residues include Leu34 and Ile58 (PmScsBα numbering); their side chains interact with each other in the hydrophobic core of both PmScsBα and EcDsbDα, so their conservation suggests a key role in folding or stability. Interestingly, the tyrosine hydroxyl of the conserved residue Tyr45 (Tyr42 in EcDsbDα) is within 3.5-3.8 Å of the active site sulfurs of Cys114 and Cys118 suggesting a potential role in catalysis or substrate binding. Two other conserved residues, Tyr87 and Pro72, are located at either end of the β-hairpin. The hydroxyl group of Tyr87 is within hydrogen-bonding distance of the main chain oxygen of Pro72 (~2.6 Å), suggesting that their conservation may be important in preserving the structure of the β-hairpin. The β2-β3 loop is also highly conserved across the ScsBα sequences, suggesting an important function.

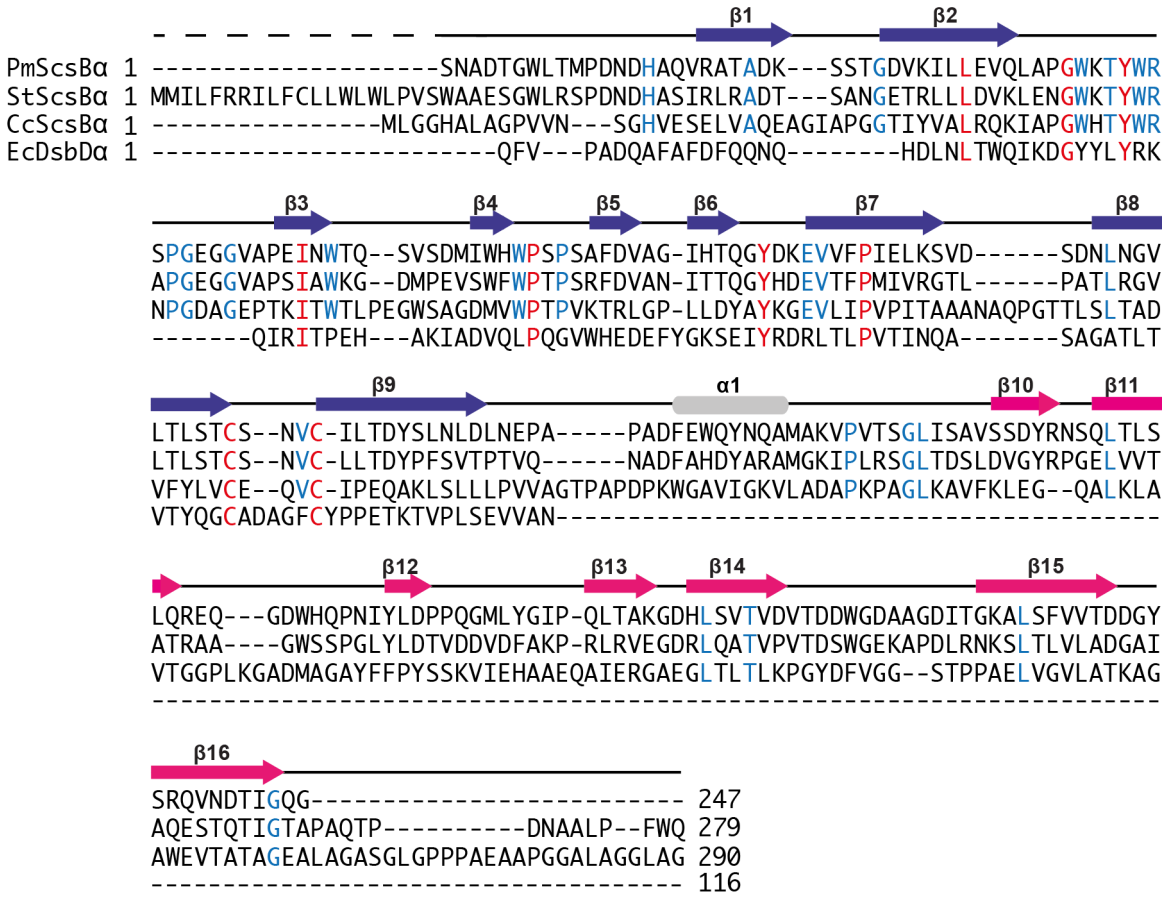


Figure 4.5: Sequence alignments. *Sequence and structure-based sequence alignment of the PmScsBα crystal structure with StScsBα, CcScsBα sequences and the EcDsbDα crystal structure (PDB: 1JPE). Secondary structure elements of PmScsBα are shown above the corresponding sequence and are coloured as per the scheme in Figure 4.3. Residues that are conserved between all four proteins are highlighted in red; those that are conserved only between the ScsBα proteins are coloured blue.*

4.4.5 Formation of a 3:1 PmScsC-PmScsBα complex

The interaction between the catalytic cysteines of PmScsC and PmScsBα is transient. For this reason the catalytic cysteine variants, PmScsC C87S and PmScsBα C114A, were produced. These variants allowed the capture of a stable PmScsC-PmScsBα complex for structural studies. Size exclusion chromatography (SEC) and subsequent SDS-PAGE analysis suggested that the predominant complex species had a 3:1 ratio of PmScsC to PmScsBα, corresponding to one PmScsC trimer bound to one PmScsBα monomer (Figure 4.6). The predicted molecular weight calculated from SAXS and SANS analysis (Table 4.2) also provided evidence to support this stoichiometry. A peak corresponding to a larger species was seen on the SEC chromatogram (peak marked Z, Figure 4.6A), suggesting that a smaller amount 3:2 and/or 3:3 complex may have also formed.

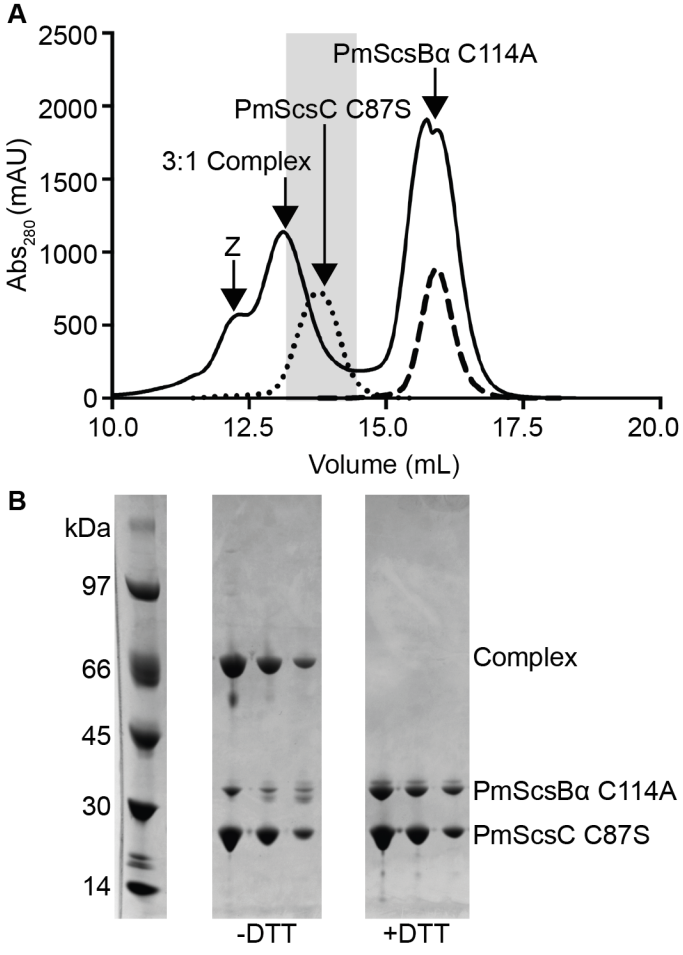


Figure 4.6: Formation of 3:1 PmScsC C87S:PmScsBα C114A complex. A. *SEC chromatograms from the purification of the 3:1 PmScsC:PmScsBα complex (solid line), purified PmScsC C87S (dotted line) and PmScsBα C114A (dashed line). In the purification of the complex, excess PmScsC was removed prior to SEC using an IMAC step. The leading shoulder (Z) suggests the presence of 3:2 and/or 3:3 PmScsC C87S:PmScsBα*

C114A complexes. The shaded region highlights the fractions of the 3:1 PmScsC:PmScsBα peak (solid line) that were pooled for structural analysis. ***B.*** *The fractions from the shaded region of the 3:1 PmScsC:PmScsBα complex peak run on SDS-PAGE, without and with DTT. PmScsC is a trimer that separates and runs as a monomer on SDS-PAGE. The lane labelled –DTT shows that the trimer:monomer complex separates into complex (covalently linked protomer of PmScsC with monomer of PmScsBα - 55.1 kDa) and uncomplexed PmScsC C87S (24.8 kDa), in an approximate 1:1 ratio. When run with DTT (lane labelled +DTT), the higher band dissociates into the two complex components in an approximate 3:1 ratio.*

4.4.6 SAXS and SANS provide structural insights into the PmScsC-PmScsBα interaction

SANS contrast variation experiments were performed on the 3:1 PmScsC-PmScsBα complex, formed using unlabelled PmScsC C87S and deuterium labelled PmScsBα C114A (^{D}PmScsBα C114A) (Figure 4.7). Deuterium labelling changes the neutron scattering length density of PmScsBα. It is possible to tune the neutron scattering length density of the solvent, by changing its deuterium content, such that it matches one of the components of the protein complex. This is said to be the match-point of that component, and the measured scattering data can then be interpreted as being from the unmatched component of the complex alone. Data collected close to the ^{D}PmScsBα C114A match-point (100% D_2O, where PmScsC C87S dominates the scattering), indicates the highly dynamic PmScsC protein becomes comparatively rigid and adopts a conformation that, at low-resolution, appears symmetrical upon interaction with PmScsBα. Specifically, the pair-distance distribution function is bimodal (Figure 4.8A, green curve) with two well defined peaks, and similar to the $p(r)$ generated from the extended symmetrical crystal structure (PDB: 5ID4) (Figure 4.8A, black dashed curve) (175). An asymmetric arrangement of the PmScsC trimer would yield an asymmetric shaped second peak in the $p(r)$, while significant flexibility would see this same peak broaden.

Analysis of the SANS data showed that a plot of $I(0)$ (the scattering intensity at zero-angle) normalised by protein concentration versus deuterium content of the solvent was well represented by a parabolic function indicating the samples were pure (Figure 4.8B). A Stuhrmann plot revealed that the ^{D}PmScsBα C114A molecule was located towards the centre of the complex (Figure 4.8C). A rigid body model of the complex was optimised simultaneously against the SAXS data set and five SANS contrast variation data sets. The

resulting model provided an excellent fit to all of the scattering data (Figure 4.9A). The configuration adopted by PmScsBα and PmScsC in the model precludes the simultaneous binding of two (or three) PmScsBα molecules to PmScsC, helping to explain the preferential formation of a trimer:monomer complex (Figure 4.6). The model also highlights two potential interfaces between PmScsC and PmScsBα (Figure 4.9A). The most obvious interface is that between chain A of PmScsC and subdomain A of PmScsBα for which the intermolecular disulfide bond is modelled. The resolution of the model is too low to be definitive, but the interface includes residues that could form hydrogen bond interactions. The second interface is between chain A of PmScsC and subdomain B (the second immunoglobulin domain) of PmScsBα.

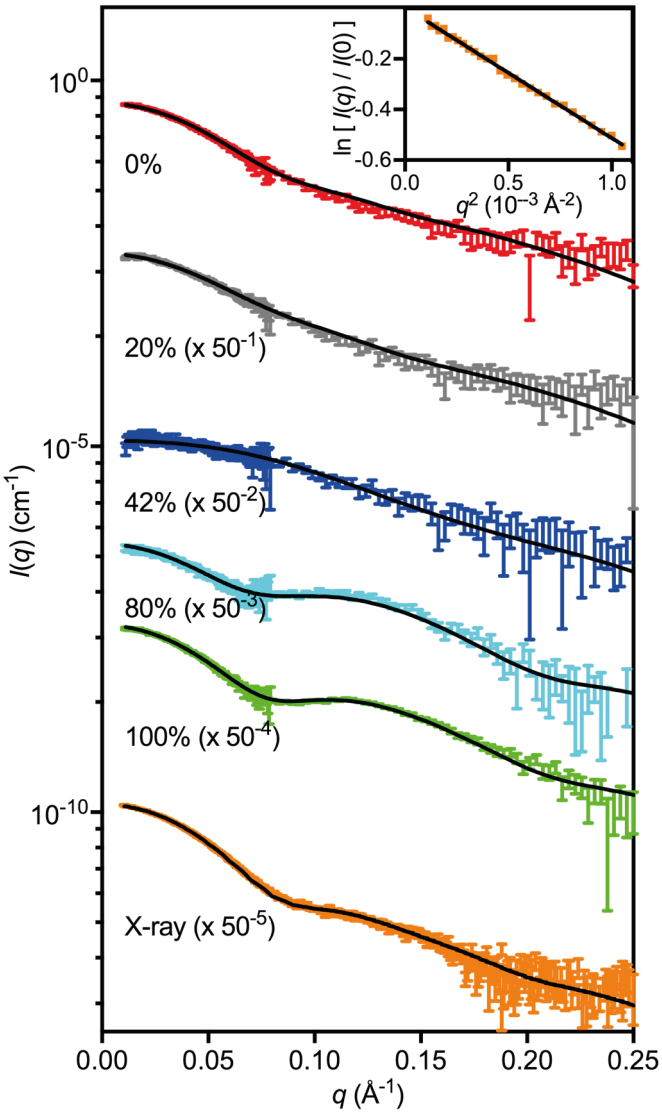


Figure 4.7: SAXS/SANS scattering curves. *SAXS and SANS data (offset for clarity) collected from the 3:1 PmScsC C87S-^{D}PmScsBα C114A complex with the model scattering curves overlaid (solid black line): 0% (χ^2 = 1.16, red, on absolute scale); 20% (χ^2 = 1.20, grey, offset by a factor of 50^{-1}); 42% (χ^2 = 1.00, blue, offset by a factor of 50^{-2}); 80% (χ^2 = 1.11, cyan, offset by a factor of 50^{-3}); 100% (χ^2 = 0.86, green, offset by a factor of 50^{-4}); X-ray (χ^2 = 1.80, orange, offset by a factor of 50^{-5}). A Guinier plot for the SAXS data is shown is shown in the inset. The match point of PmScsC C87S is ~42% D_2O (blue) where ^{D}PmScsBα C114A dominates the scattering, while the match point of ^{D}PmScsBα C114A is ~100% D_2O (green) where PmScsC C87S dominates the scattering.*

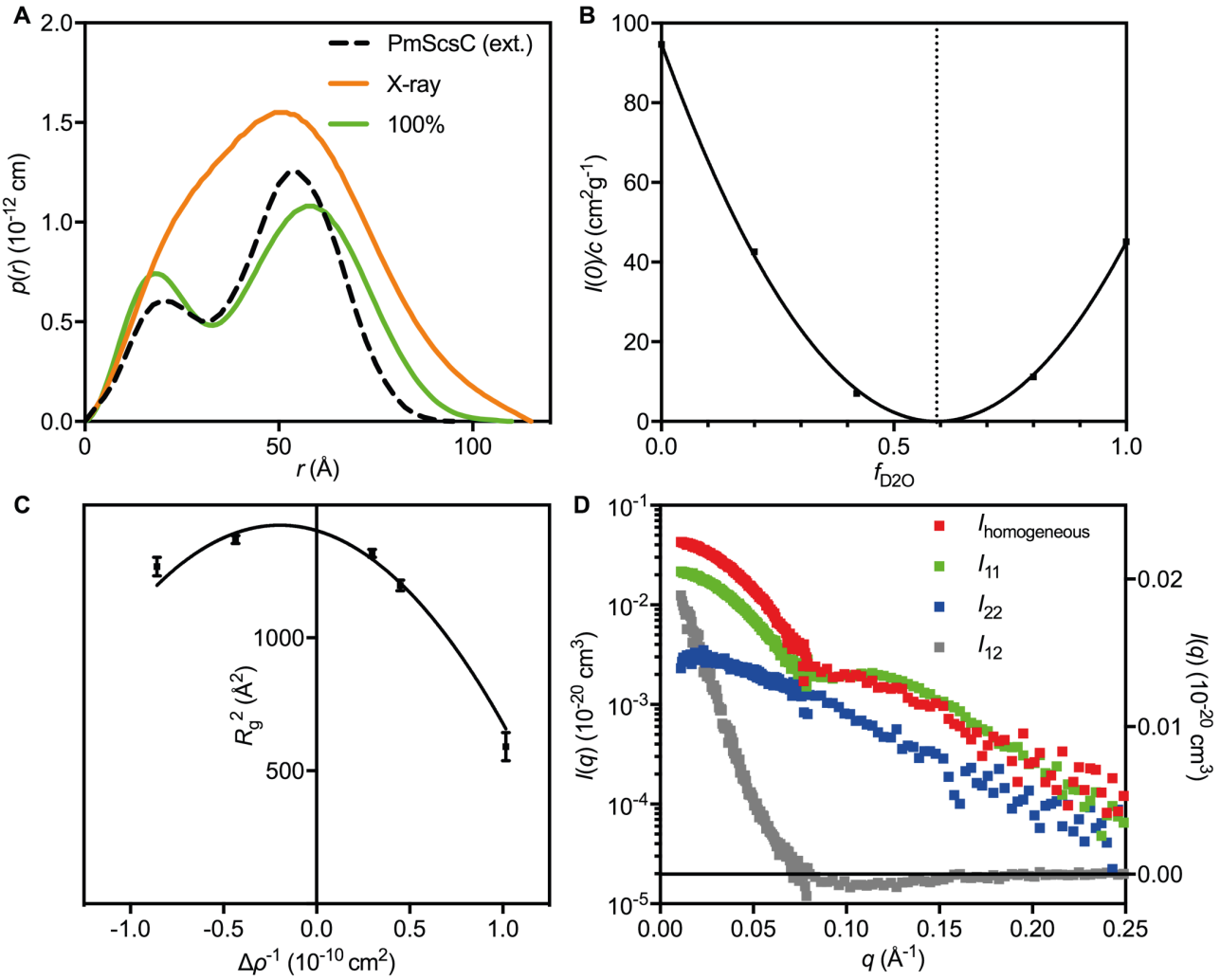


Figure 4.8: Other SAXS/SANS results. A. *Pair-distance distribution profile, p(r), derived from the experimental X-ray scattering data (orange), and the 100% scattering data (green), which should approximate the p(r) for PmScsC C87S alone. The p(r) derived from the extended crystal structure (PDBID: 5ID4) is shown as a black dotted line for reference.* ***B.*** *A plot of I(0) normalised by concentration as a function of D_2O content of the supporting solvent. The plot is parabolic in shape and reveals that the match-point of the entire complex is 59% D_2O (vertical dotted line).* ***C.*** *Stuhrmann plot for the 3:1 PmScsC C87S-^{D}PmScsBα C114A complex, conforming to $R_g^2 = R_m^2 + \alpha\Delta\rho^{-1} - \beta\Delta\rho^{-2}$, where R_m is the radius of gyration of that object with homogenous contrast, and α and β are related to the contrast variations within the object. The values obtained from a fit to the plot (solid black line) are: $R_m = 37.5 \pm 0.4$; $\alpha = -210 \pm 50$; and $\beta = 520 \pm 100$. The negative value of α reveals that the region with higher contrast (i.e. ^{D}PmScsBα C114A) lies towards the centre of the molecule.* ***D.*** *The composite scattering functions determined from the neutron contrast variation data. The I_{11} curve (green, left y-axis) corresponds to scattering from PmScsC C87S, the I_{22} curve (blue, left y-axis) corresponds to scattering from ^{D}PmScsBα C114A, and the I_{12} curve (grey, right y-axis) is related to the arrangement of ^{D}PmScsBα*

C114A relative to PmScsC C87S. $I_{homogeneous} = I_{11} + I_{22} + I_{12}$ (red, left y-axis) is the scattering curve of an object with the same shape as the PmScsC C87S-^{D}PmScsBα C114A complex, but with homogeneous contrast, and is used for estimating the Porod volume and molecular mass from the SANS data.

Table 4.2: SAS data collection and analysis details for the PmScsC-PmScsBα complex.

Data Collection Parameters	**SAXS**	**SANS**
Instrument	SAXS-WAXS (Australian Synchrotron)	QUOKKA (ANSTO)
Beam geometry	Point	Point
Wavelength (Å)	1.033	5.00
Sample to detector distance (m)	2.680	2.000 (short) 10.000 (long)
q-range (Å^{-1})	0.007 – 0.35	0.03 – 0.40 (short) 0.01 – 0.09 (long)
Exposure time (s)	13 (13 × 1 s exposures)	0%^{S}: 7200; 20%^{S}: 10800; 42%^{S}: 21600; 80%^{S}: 7200; 100%^{S}: 3600; 0%^{B}: 10800; 100%^{B}: 7200
Measurement type	Concentration series from 96-well plate, 1.0 mm quartz capillary	Neutron contrast variation, Hellma 120-QS 1.0 mm quartz cells
Protein concentration range (mg/mL)	0.56 – 4.50	-
Temperature (K)	283	289
Absolute intensity calibration	Water	Incident beam intensity
Sample details		
Extinction coefficient (A_{280}, 0.1% *w/v*)	0.979	0.979
Partial specific volume (cm^{3}g^{-1})	0.737	0.737
Contrast, $\Delta\rho$ (10^{10} cm^{-2})	2.86	3.36 (0%); 2.23 (20%); 1.10 (42%); -1.17 (80%); -2.30 (100%)
Molecular mass [from sequence] (kDa)	105.5	105.5
Protein concentration (mg/mL)*	0.56	5.30 (0%^{S}); 5.20 (20%^{S}); 5.10 (42%^{S}); 5.00 (80%^{S}); 4.90 (100%^{S})
Structural parameters		
I(0) (cm^{-1}) [from Guinier]	0.04060 ± 0.00010	-
R_g (Å) [from Guinier]	39.2 ± 0.2	-
I(0) (cm^{-1}) [from $p(r)$]	0.04026 ± 0.00010	0.5020 ± 0.0019 (0%) 0.2217 ± 0.0013 (20%) 0.0357 ± 0.0007 (42%) 0.0561 ± 0.0008 (80%)

		0.2211 ± 0.0010 (100%)
R_g (Å) [from $p(r)$]	38.7 ± 0.1	36.3 ± 0.2 (0%)
		34.6 ± 0.3 (20%)
		24.3 ± 1.1 (42%)
		35.6 ± 0.5 (80%)
		37.0 ± 0.2 (100%)
D_{max} (Å)	115 ± 3	110 ± 3
Porod volume (Å^3)	131300 ± 7000	116000 ± 6000†
Molecular mass determination		
Molecular mass [from $I(0)$] (kDa)	97 ± 5	95 ± 5†
Molecular mass [from Porod] (kDa)	108 ± 5	95 ± 5

* $I(0)$ = 0.0854 ± 0.0002 cm^{-1}, R_g = 39.0 ± 0.1 Å, M = 103 ± 5 kDa (1.13 mg/mL); $I(0)$ = 0.1779 ± 0.0002 cm^{-1}, R_g = 39.0 ± 0.1 Å, M = 107 ± 5 kDa (2.25 mg/mL); $I(0)$ = 0.3476 ± 0.0003 cm^{-1}, R_g = 38.8 ± 0.1 Å, M = 105 ± 5 kDa (4.50 mg/mL). There is a small systematic change in R_g but no systematic change in M, thus it was deemed that the complex is largely free of concentration dependent effects up to a concentration of ~5.0 mg/mL. † The Porod volume calculation is not valid for systems where the $\Delta\rho$ of the particle is not homogeneous (such as is the case for the neutron contrast variation experiment reported here). Instead, the composite scattering functions (Figure 4.8D) were summed together to give the scattering curve of the protein complex with homogeneous contrast (i.e. $I_{homogeneous} = I_{11} + I_{22} + I_{12}$), and the Porod volume was determined from this curve. The molecular mass derived from the $I(0)$ was also taken from this curve.

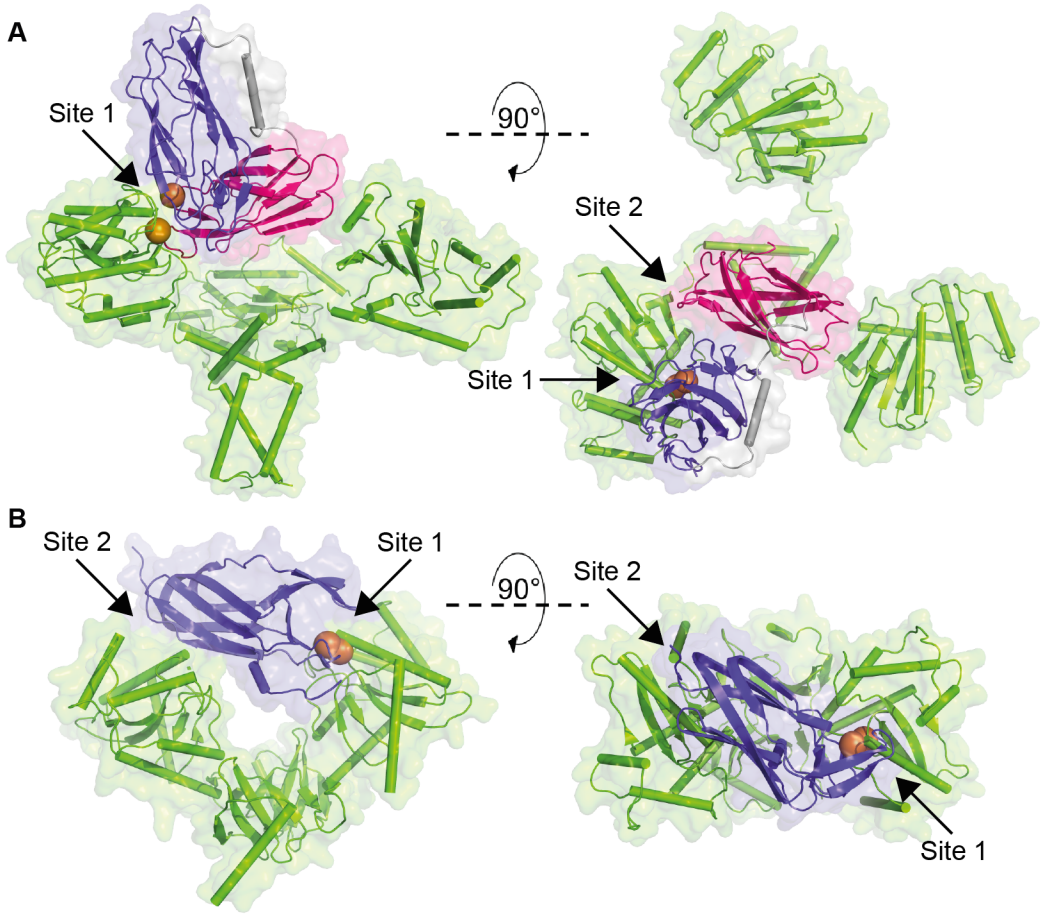


Figure 4.9: Binding sites in the PmScsC-PmScsBα SANS model and EcDsbC-EcDsbDα structure (PDB: 1JZD). A. *Model of the 3:1 complex generated from the SAXS/SANS data, shown in backbone and surface representation. PmScsBα binds to one PmScsC (green) protomer with both subdomain A (purple, site 1) and B (magenta, site 2).* ***B.*** *Crystal structure of the EcDsbC-EcDbsDα complex. EcDsbDα (purple) binds to both protomers of EcDsbC (green, binding sites 1 and 2). In both A and B, the catalytic cysteines involved in the intermolecular disulfide bond between the partner proteins are shown as orange spheres.*

4.5 Discussion

Suppressor of copper sensitivity proteins play a role in the response to copper stress of the important human pathogens *S.* Typhimurium and *P. mirabilis* (120,175). At least two of the four Scs proteins are homologues of Dsb proteins (111), which are essential for the correct formation of disulfide bonds in virulence factors of many Gram-negative bacteria (16,30,193,194). *P. mirabilis* ScsC was recently shown to be a trimeric, highly dynamic disulfide isomerase (175). The present study aimed to characterise PmScsBα - the predicted interaction partner of PmScsC - and investigate the nature of the interaction between the two.

The domain architecture of full-length ScsB resembles that of *E. coli* DsbD, the major difference being the size of the periplasmic N-terminal domain. This difference suggested that ScsBα would be structurally distinct from DsbDα perhaps because they interact with different substrates. We confirmed that PmScsC and PmScsBα are redox partners, and are therefore functionally similar to EcDsbC and EcDsbDα. We also reported the structure of PmScsBα, showing that it consists of two immunoglobulin-like folds connected by an α-helix. The structures of the two immunoglobulin-like subdomains of PmScsBα are broadly similar with the exception that subdomain A has a catalytic CXXXC motif, as well as a β-hairpin and extended loop that both cover the CXXXC active site. The positioning of the β-hairpin and the loop could potentially shield the active site from non-specific interactions or facilitate interactions with partner proteins. Unexpectedly, we found that domain A of PmScsBα shares the same features of the immunoglobulin fold of EcDsbDα (Figure 4.4, RMSD: 2.5 Å, 103 residues aligned) - despite a sequence identity of just 12%. The key difference between the structures of subdomain A and EcDsbDα is that EcDsbDα lacks the β2-β3 loop of PmScsBα. This β2-β3 loop is conserved in *S.* Typhimurium and *C. crescentus* ScsBα suggesting that it plays a key functional role. A residue that is conserved across the three ScsBαs and EcDsbDα is Tyr45 (PmScsBα numbering). In EcDsbDα, the equivalent Tyr42 hydroxyl is thought to be involved in nucleophilic attack of Cys103, which resolves the intermolecular disulfide formed between EcDsbDα and its substrate (*e.g.* EcDsbC) (61). The nucleophilic attack in EcDsbDα is proposed to also require Asp68 (61), a residue that is not conserved in Pm, Cc or StScsBα. It is unclear from the PmScsBα structure what residue, if any, might play an equivalent role to Asp68.

The discovery of tandem immunoglobulin-like folds in PmScsBα raises the question as to the function of the second or non-catalytic immunoglobulin-like fold. EcDsbDα has a single

immunoglobulin-like fold, which, in the *E. coli* DsbC-DsbDα crystal structure, spans the central cleft of the V-shaped DsbC substrate to interact with both domains of the disulfide isomerase (Figure 4.9B) (47). This binding mode is thought to favour binding to dimeric EcDsbC; EcDsbDα is unable to interact with monomeric EcDsbC (44,47). By comparison, the SANS-derived model of the PmScsC-PmScsBα complex suggests that PmScsBα subdomain A interacts with just one of the three PmScsC protomers in the trimer. Subdomain B of PmScsBα forms a secondary binding site to PmScsC that may stabilise the primary interaction or add specificity (Figure 4.9A). Subdomain B could also be required for interactions with other proteins that have yet to be identified.

EcDsbC-EcDsbDα is the redox relay complex formed between dimeric EcDsbC and monomeric EcDsbDα. The crystal structure revealed a 2:1 (rather than a 2:2) complex between the dimer of EcDsbC and a monomer of EcDsbDα (44,47). On this basis, we hypothesised that PmScsC-PmScsBα would form a 3:1 complex between trimeric PmScsC and monomeric PmScsBα, and the SANS structure of the complex appears to preclude higher stoichiometry. Nevertheless, we did find *in vitro* evidence suggestive of PmScsC forming 3:2 and 3:3 complexes with PmScsBα (Figure 4.6A), showing that different arrangements might form with PmScsBα in the absence of the PmScsB β and γ domains. We expect that these higher order complexes would not form in the context of full-length PmScsB.

How do our findings on PmScsBα and PmScsC inform on interactions between ScsC and ScsB in other organisms? A functional interaction between ScsC and ScsB has already been reported for *C. crescentus* (118). With two examples now confirmed, the ScsB-C redox relay system is likely to be present in all organisms that encode the 4-gene *scs* cassette. However, there is likely to be a wide diversity of interactions. ScsC is reported to be monomeric for StScsC (120), dimeric for CcScsC (118), and trimeric for PmScsC (175). By contrast, ScsB shares similar features across the same organisms suggesting that StScsBα and CcScsBα both incorporate the tandem immunoglobulin-fold arrangement present in PmScsBα. As indicated above, both subdomains of PmScsBα interact with one PmScsC protomer, suggesting that this tandem immunoglobulin fold arrangement could facilitate the interaction with ScsCs regardless of the ScsC oligomeric state. Future studies will be needed to establish whether it is the tandem immunoglobulin-fold architecture that permits interactions with monomeric, dimeric and trimeric PmScsCs. Curiously, a BLAST search of *E. coli* DsbD or *S.* Typhimurium DsbD against the *P. mirabilis* genome did not

identify any hits, suggesting the possibility that PmScsB may also act as the redox partner for *P. mirabilis* DsbC proteins.

4.6 Experimental procedures

4.6.1 Sequence analysis

The TMHMM Server v. 2.0 was used to determine the region of the ScsB proteins that formed the N-terminal periplasmic domain. EMBOSS Needle (195,196) with residues 1-275, 1-279 and 1-290 of UniProt entries B4EV20 (PmScsB), AAL20046 (StScsB) and Q9ABL0 (CcScsB), respectively, was used to calculate the sequence identities reported between these proteins. The DALI sever (192) was used to calculate RMSDs and structure based sequence identities. Coot (156) was used to create visuals of the structural alignments. The sequences of all four proteins were aligned using PROMALS3D (197) with the PDB files of PmScsBα and EcDsbDα (1JPE) instead of the sequence alone.

Residue numbers quoted in the text were taken from the deposited PDB files for the PmScsC and PmScsBα protein structures, which are numbered based on the TEV cleaved protein sequence (starting $S^1N^2A^3$…).

4.6.2 Molecular Biology

A codon-optimised gene encoding *P. mirabilis* ScsBα without the predicted signal peptide (UniProt ID: B4EV20, residues 21-275, Integrated DNA Technologies) was inserted into pMCSG7 (198) using ligation independent cloning. The PmScsC (UniProt ID: C2LPE2/ B4EV21) construct used throughout this study was created previously (175). The constructs for the variants PmScsC C87S and PmScsBα C114A were created using the QuikChange® Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). The wild-type PmScsC and PmScsBα constructs were used as templates and primers 5'-gcggatcgaaacgtttagagtacgggcaattatag-3', 5'-ctataattgcccgtactctaaacgtttcgatccgc-3' and 5'-gtcagaatgcatacattgctagcggtactcaatgtaaggacacc-3',
5'-ggtgtccttacattgagtaccgctagcaatgtatgcattctgac-3', respectively, were used to introduce the point mutations. PmScsBα was also subcloned into pET24a for deuterated bioreactor expression, using restriction enzymes *Eco*RI and *Nde*I. All constructs contained an N-terminal His_6 tag followed by a TEV protease cleavage site and were confirmed by sequencing.

4.6.3 Protein production

Except for the selenomethionine labelled PmScsBα and deuterated PmScsBα C114A, all proteins were expressed and purified as outlined previously (175). The redox state of the cysteines was confirmed spectrophotometrically with 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB or Ellman's reagent) as described previously (199). Yields for PmScsC and PmScsBα derivatives were around 80 mg and 40 mg per litre of culture (~15 g cell weight harvested), respectively. SDS-PAGE (NuPAGE® system, 4-12% BisTris gel, Invitrogen, Australia) with Coomassie blue stain was used to assess the protein quality.

4.6.4 Expression of SeMet labelled PmScsBα

To express selenomethionine (SeMet) labelled PmScsBα, needed for experimental phasing in crystallographic structure determination, a 10 mL starter culture of *E. coli* BL21 (DE3) pLysS cells containing pMCSG7-PmScsBα was grown in LB media supplemented with the appropriate antibiotics at 37°C for 4-5 h, 220 rpm. This culture was then used to inoculate 1L of M9 minimal media supplemented with antibiotics and amino acids (100 mg of L-methionine, L-lysine, L-threonine, L-phenylalanine and 50 mg of L-leucine, L-isoleucine, L-valine), which was incubated overnight at 37°C, 220 rpm. Cells were harvested from the overnight culture, washed with 1L of sterile water and resuspended in water to an OD_{600} of ~1. M9 minimal media containing antibiotics but no amino acids was then inoculated with the cells to a final OD_{600} of 0.1. The cells were cultured at 30°C, 220 rpm for 6-8 h, until an OD_{600} of 0.5 was reached. Amino acids, including 100 mg of L-selenomethionine instead of L-methionine, were then added to the culture flasks and after 15 min at 30°C, IPTG was added to a final concentration of 1 mM to induce expression. The cultures were incubated at 30°C, 220 rpm overnight. Cells were harvested and the protein was purified as described previously (175) with the addition of 5 mM DTT to all purification buffers. Yield of SeMet labelled PmScsBα was 5 mg per litre of culture (~5 g of cell pellet).

4.6.5 Expression of Deuterated PmScsBα C114A

Deuterated PmScsBα C114A was expressed for the small-angle neutron scattering (SANS) analysis of the complex between PmScsC C87S and PmScsBα C114A. The deuterated protein PmScsBα C114A was produced at the National Deuteration Facility, Australian Nuclear Science and Technology Organisation in 1 L "ModC1" media (200) containing 90% D_2O *v/v* and unlabelled glycerol (40 g/L). Miniprep plasmid DNA "pET24a-PmScsBα C114A" was used to transform 50 µL of Invitrogen competent BL21*(DE3) cells,

which were then incubated with 250 µL of SOC media for 2 h. The 300 µL culture was added to 10 mL of ModC1 media containing 50% D_2O, and 40 µg/mL kanamycin, shaking at 200 rpm, 37 °C. After 16 h (OD600, 1.2), 9 mL was added to 36 mL fresh media with kanamycin at 100% D_2O (i.e. 45 mL at 90% D2O). After another 5 h (OD600, 0.9), the culture was diluted with fresh 90% D_2O media to 102 ml. After another 5 hours, 100 mL (OD600, 0.87) was used to inoculate 900 mL of fresh media 90% D_2O with kanamycin in a 2L Real Time Engineering bioreactor, aerated with air at 0.5 L/min, and with pH held to a minimum of 6.2 by addition of 28% NH_4OH in H_2O (Sigma). Dissolved oxygen tension was set to 75% and was above 60% throughout. At OD600 = 14.2, the temperature was lowered to 20°C and expression was induced by the addition of 1 mM IPTG. At 17 h post-induction, with OD600 = 32, the culture was harvested with a wet weight yield of 79.4 g, and with excellent expression evident by SDS-PAGE. The deuteration level was determined by partial trypsin digest MALDI-TOF comparison of unlabelled and labelled samples and was found to be 70.8%. Deuterated PmScsBα C114A was purified as described for the unlabelled protein, except the His_6-tag was not cleaved, as it was needed for the purification of the complex. The yield of purified deuterated PmScsBα was ~240 mg per one litre of bioreactor culture (~40 g of cell pellet).

4.6.6 Determination of protein concentration

The reported protein concentrations were determined using the A_{280} of the sample (read using a Thermoscientific NanoDrop 2000c spectrophotometer) and calculated extinction coefficients from ProtParam (140).

4.6.7 Crystal structure of PmScsBα

A number of conditions from commercial screens resulted in PmScsBα crystals. These conditions were replicated in 24-well plates and the best diffracting SeMet labelled PmScsBα crystals were obtained in a drop containing 1 µL 43 mg/ml protein in 10 mM HEPES pH 7.4, 5 mM DTT and 1 µL of 0.1 M MES monohydrate pH 6, 6% v/v tacsimate pH 6 and 24% PEG4000. A commercial additive screen (Hampton Research) was used to optimise crystals and those grown over 2 days to ~400 µm with the addition of 4% 2,5-hexanediol or 4% 1,3-butanediol resulted in the best diffraction. Crystals were cryoprotected in the crystallisation condition plus 20% glycerol and flash frozen in liquid nitrogen.

Due to the lack of a suitable model for molecular replacement, experimental phasing was required to solve the crystal structure of PmScsBα. A fluorescence scan was used to

determine the absorption edge of SeMet labelled crystals and a data set (Table 4.1) was collected near the peak absorption wavelength (0.9786) at 100 K. Data processing in autoPROC (159), which utilises XDS (151) and AIMLESS (152), determined the space group to be $P2_12_12_1$. SAD phasing using the SHELX workflow (201) in HKL2MAP (202) found 14 Se sites with an occupancy > 0.2. We expected only 12 Se sites - the extra two sites were positioned very close to other sites, suggesting alternate conformations, though only one SeMet was modelled that way in the final structure. SAD phasing resulted in an initial incomplete alanine model of the protein structure (FOM 0.42). This initial model was then used for molecular replacement against the same data in Phaser (155) in the CCP4 suite (160). Two rounds of automated building and refinement using BUCCANEER (161) with REFMAC (158), as implemented in CCP4 (160), completed the structure, revealing three molecules in the asymmetric unit. Then several rounds of manual adjustment in COOT (156) and refinement in PHENIX (154) were performed, with reference to the validation program MolProbity (164). Multiple residues in the crystal structure have alternate conformers and occupancies of the selenium atoms in the SeMet residues were all <1.0. The electron density for some loop regions is poor in chains B and C, particularly residues 79-82, 99-100 and 51-52 (Chain B only). R/Rfree values are 17.3% and 19.8% at a resolution of 1.54 Å. All structural figures were made in PyMOL (203).

4.6.8 Redox interaction assay

To confirm the redox interaction between PmScsC and PmScsBα, oxidised PmScsC and reduced PmScsBα were incubated together at final concentrations of 10 μM in 100 mM sodium phosphate pH 7, 1 mM EDTA for 2 h. As controls, oxidised and reduced forms of both proteins, a sample containing reduced PmScsC and oxidised PmScsBα and a sample with oxidised PmScsC with PmScsBα C114A were also incubated for 2 h. All samples were then treated with 10% TCA to precipitate the protein, and after centrifugation, the pellets were washed with 100% cold acetone and dried. The protein pellets were resuspended in a solution of 200 μM 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS) in 1% SDS, 50 mM Tris pH 7.0. AMS binds to the reduced cysteines of proteins, increasing the molecular weight and allowing the differentiation between the oxidised and reduced forms of a protein when run on a gel. Therefore, all samples were run on a non-reducing, non-denaturing 12% NuPAGE Bis-Tris gel for 2.5 h at 100V. The gel was Coomassie stained and imaged using an Odyssey® Fc Imaging system (LI-COR® Biosciences) with the 700 nm channel.

4.6.9 RNaseA isomerase assay

To determine whether PmScsBα activates the isomerase PmScsC, an isomerase assay utilising scrambled RNaseA (scRNaseA) and the RNase A model substrate, cytidine 2':3'-cyclic monophosphate (cCMP) was performed. Bovine RNase A (purchased from Sigma-Aldrich) was scrambled as previously described (170). The various oxidised and reduced protein samples described above for the gel-shift redox interaction assay were again prepared but with the addition of 40 μM scRNase A to each sample. After 2 h, 50 μL of the protein/scRNase A reactions was taken and added to 150 μL of 4 mM cCMP in a UV-star 96-well plate (Greiner Bio-One, Austria). The increase in absorbance at 296 nm was monitored over 3 min with a Synergy H1 multimode plate reader (BioTek, USA) and used to calculate the rate of cCMP hydrolysis. Native RNase A and scRNase A only samples were included as controls and used to determine the relative activity of the RNase A in each sample. The results shown are the average and standard deviation of 3 technical replicates.

4.6.10 Complex formation

The PmScsC and PmScsBα cysteine variants were designed following the same approach used for the EcDsbC:EcDsbDα interaction (47). Mutation of cysteines in EcDsbC and EcDsbDα resulted in a stable complex between the two proteins (44). Cysteines C87 of PmScsC and C114 of PmScsBα are the least accessible of the two cysteines in each active site, suggesting they facilitate resolution of intermolecular disulfide bonds. Variants PmScsC C87S and PmScsBα C114A were therefore prepared and used in complex formation studies.

The 3:1 PmScsC C87S:PmScsBα C114A complex was formed by mixing the two proteins together in a 3:1 ratio at concentrations of approximately 200 mg/ml and incubating the mixture at 37°C for 3 days. To separate the complex and free PmScsBα from PmScsC, PmScsBα was left His-tagged and the protein mixture was then run over 2 mL of fresh TALON resin, twice. Columns were washed with 25 mM Tris pH 7.4, 150mM NaCl and the complex and free PmScsBα were eluted with 25 mM Tris pH 7.4, 150mM NaCl, 250 mM imidazole. This was then run over a Superdex 200 10/300 Increase size exclusion column on an ÅKTA system to separate PmScsBα from complex. A peak on SEC eluted earlier than for ScsC, suggesting that a complex between the two proteins had formed (Figure 4.6A). Running the fractions from the peak corresponding to the 3:1 complex on non-reducing SDS-PAGE resulted in three bands (Figure 4.6B). The lower bands correlated to free PmScsC C87S (24.8 kDa) and a small amount of PmScsBα C114A (30.3 kDa) and

the upper band corresponded to the molecular weight of one PmScsC C87S molecule bound to one PmScsBα C114A molecule (55.1 kDa). The upper band dissociated when reducing agent was added, leaving two bands correlating to monomeric PmScsC C87S and PmScsBα C114A, in an approximate 3:1 ratio (Figure 4.6B). There was also evidence for the formation of small amounts of complex containing PmScsC C87S and PmScsBα C114A in ratios of 3:3 and 3:2 but only those SEC fractions containing 3:1 complex were collected for further analysis. The complex formation process was repeated multiple times to gain enough sample for SAXS/SANS analysis as the overall yield of 3:1 complex was ~2.5%.

4.6.11 Small-angle solution scattering

Small-angle neutron scattering data shown in Table 4.2 were collected on the QUOKKA instrument at the Australian Nuclear Science and Technology Organisation (Table 4.2) (204). Two aliquots of PmScsC C87S-^{D}PmScsBα C114A complex at ~5 mg/mL were dialysed – one against a buffer with a D_2O content of 0% and the other against a buffer with a D_2O content of ~100%. Samples with D_2O contents of 20%, 42%, and 80% were obtained by mixing the two original solutions and the appropriate ratios. The two-dimensional data were normalised to a common incident neutron count and corrected for sample transmission, background radiation, empty cell scattering, detector sensitivity and radially averaged to produce $I(q)$ vs. q profiles. Scattering data from the two different sample-detector distances were then merged and buffer scattering data were then subtracted from the protein + buffer data to give the protein scattering profiles (the 20%, 42%, and 80% buffer scattering curves were taken as a linear combination of the 0% and 100% buffer scattering curves). To correct for the effects of incoherent scattering by ^{1}H-rich samples, backgrounds levels were adjusted by a small constant such that the high-q scattering displayed q^{-4} dependence.

Small-angle X-ray scattering data were collected on the SAXS-WAXS beamline at the Australian Synchrotron (Table 4.2) (142). Serial dilutions of the stock solution of PmScsC C87S-^{D}PmScsBα C114A complex were made to give four concentrations between 0.55 and 4.5 mg/mL that were centrifuged at 10,000 × g immediately before loading into a 96-well plate. To reduce radiation damage, samples (~100 µL) were drawn into a capillary from the 96-well plate and flowed past the beam. All measured two-dimensional data were averaged and corrected for sample transmission, solvent scattering, detector sensitivity, and radially averaged to produce I(q) vs. q profiles using Scatterbrain (v 2.7.1) (143).

The estimated molecular masses were calculated using values for contrast and partial specific volume predicted from the protein sequence using *MULCh* (v 1.1) (145). Data processing and Guinier analysis was performed using Primus (v 3.2) (185). The pair-distance distribution function, $p(r)$, was generated from the experimental data using *GNOM* (v 4.6) (4), from which $I(0)$, R_g and D_{max} were determined. For the SANS data, analysis of the dependence of R_g upon contrast, and calculation of the composite scattering functions, was also calculated using *MULCh* (Figure 4.8D). Co-refinement of a rigid-body model against SANS and SAXS data was performed using SASREF (v 7.d) (205) where the crystal structure of PmScsC (PDB: 4XVW – Chain A) was broken into three rigid sections: 1) residues A3 – Q43 that were fixed as to preserve the structure of the trimerisation domain observed in the crystal structure; 2) residues Q44 – F47 were defined as a short linker with a restraint between residues Q43 and Q44 to keep them in close proximity; 3) residues R48 – K224 were defined as the catalytic domain free to move with a restraint between residues F47 and R48 to keep them in close proximity. The additional PmScsC chains were generated by symmetry (assuming a C_3 point group). The crystal structure of PmScsBα (residues A2 – A248) was included in the optimisation with a distance restraint between C84 of a PmScsC monomer and C118 of PmScsBα of 8 Å. Thus, while the complex overall is not symmetrical, the PmScsC portion of the model possesses 3-fold symmetry. The optimisation was run 16 times and the model that provided the best global fit to the scattering data is reported here, however, all models obtained are similar in structure.

4.7 Footnotes

The atomic coordinates and structure factors of PmScsBα have been deposited in the Protein Data Bank under the ID # 6C29. SAXS/SANS data and model of the 3:1 PmScsC-PmScsBα complex have been submitted to SASBDB under the ID # SASDC48.

Conflict of Interest: The authors declare that they have no conflicts of interest with the contents of this article

4.8 Acknowledgements

We acknowledge the support of the Australian Centre for Neutron Scattering, the National Deuteration Facility, and the SAXS-WAXS and MX beamlines at the Australian Synchrotron (Australian Nuclear Science and Technology Organisation), in providing the facilities used in this work. We thank the staff and facilities of the University of Queensland

Remote Operation Crystallisation and X-ray (UQ ROCX) facility for their support. We also acknowledge access to the University of Sydney Mass Spectrometry Core Facility, which was used to produce partial trypsin digest MALDI-TOF spectra. This work was supported by an Australian Research Council Laureate Fellowship (FL0992138, J.L.M.), an Australian Research Council Discovery Project to J.L.M. (DP130100576, H.G.C. salary), an Australian Postgraduate Award (E.J.F.) and an Institute for Molecular Bioscience Research Advancement Award (E.J.F.). We would like to thank Dr Róisín M. McMahon and Signe Christensen for helpful discussions.

Chapter 5

5. Expression and Purification of PmScsB

5.1 Introduction

Thus far this thesis has described the characterisation of the trimeric disulfide isomerase PmScsC, the structure of PmScsBα and the interaction between these two proteins. The next logical step is the investigation of the full-length transmembrane protein, PmScsB, including the PmScsBβ and γ domains. ScsB was found to be essential for growth in *Caulobacter crescentus* and the same study reported three substrates of CcScsB (118). No biochemical or structural studies of the whole transmembrane protein have been performed, though bioinformatic analyses identified that ScsB proteins form a class within the DsbD superfamily (118).

5.1.1 DsbD

DsbD from *E. coli* has been well characterised and functions as a transmembrane electron transporter, providing reducing power to the oxidative environment of the periplasm (56). DsbD consists of three domains, DsbDα, DsbDβ and DsbDγ and each domain contains two redox-active cysteines that are essential for its electron transporting activity (Figure 5.1) (57,58). DsbDα is the N-terminal periplasmic domain and has an immunoglobulin-like fold (47,61). DsbDβ is the transmembrane region that consists of 8 predicted transmembrane helices, with the two catalytic cysteines on transmembrane helices 1 and 4 (58). DsbDγ has a thioredoxin fold and is the C-terminal periplasmic domain (63).

E. coli DsbD (EcDsbD) is embedded in the inner membrane and transports electrons from the cytoplasm to the periplasm through a disulfide exchange cascade (59,191) (Figure 5.1). Unlike other electron transporters it does not rely on a cofactor for activity (206). The disulfide exchange cascade begins in the cytoplasm with NADPH-dependent thioredoxin reductase reducing thioredoxin, which in turn reduces the catalytic cysteines in EcDsbDβ (48). Transmembrane helices 1 and 4 of EcDsbDβ are suggested to be in an inverted pseudo-symmetrical arrangement that forms an hourglass shaped cavity enabling the catalytic cysteines to be accessed from either side of the membrane (Figure 5.1 inset) (65,66). This arrangement allows thioredoxin to reduce the catalytic cysteines from the cytoplasmic side of the membrane and the reducing power to be passed onto the mobile arm-like domain EcDsbDγ (63) on the periplasmic side of the membrane (59,65,66). EcDsbDγ then reduces EcDsbDα, which in turn interacts with different substrates to reduce their cysteines (59). The known substrates of EcDsbD include, EcDsbC, EcDsbG

and EcCcmG (44,56,59,68), all of which have thioredoxin folds each with two catalytic cysteines that need to be reduced for activity (39,70,124). EcDsbC is responsible for the isomerisation of disulfide bonds in periplasmic proteins (53), EcDsbG protects single cysteines of periplasmic proteins from oxidation to sulfenic acid (67) and EcCcmG is involved in cytochrome c biogenesis (69).

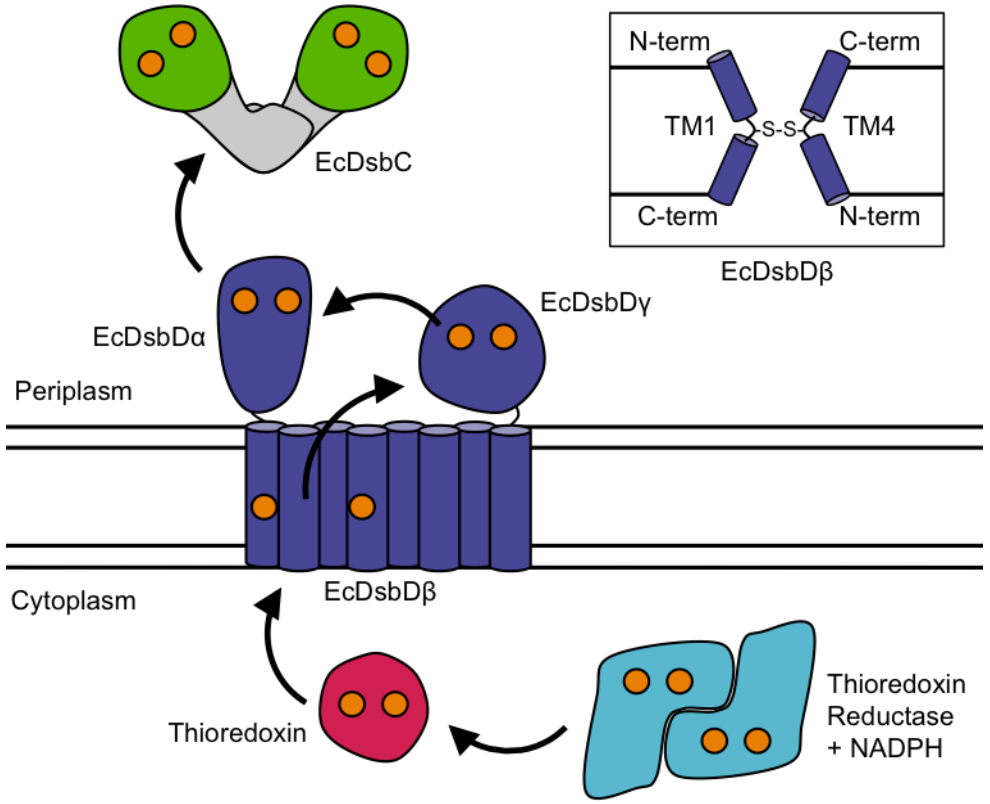


Figure 5.1: EcDsbD structure and function. *The schematic depicts the disulfide exchange cascade that allows the transfer of electrons from the cytoplasm to the periplasm. The inset figure depicts the inverted pseudo-symmetrical arrangement of transmembrane (TM) helices 1 and 4.*

To date, crystal structures have been solved of the α- (PDB ID: 1JPE/1L6P) (47,61) and γ-domains (PDB: 1UC7) (63) of EcDsbD, the complex between EcDsbDα and -γ (PDB ID: 1VRS) (64) and the complex between EcDsbDα and substrates (PDB: 1JZD, 1Z5Y) (47,125) (Figure 5.2 A-E). The structure of the full-length transmembrane protein, however, remains elusive. Two NMR structures of the EcDsbD homologue, CcdA (207), from *Archaeoglobus fulgidus* (PDB ID: 2N4X) (208) and *Thermus thermophilus* (PDB ID: 5VKV) (209) have been solved (Figure 5.2 F,G). CcdA is the simplest form of the electron transporter, consisting only of 6 transmembrane helices, whereas EcDsbDβ is predicted to have 8 transmembrane helices. The two NMR structures that have been solved are

substantially different from each other, despite high sequence similarity (29% identity, 60% similarity) (208,209) raising questions as to how applicable the findings are to EcDsbDβ. The *T. thermophilus* structure, however, is more consistent with the results from solvent accessibility assays involving EcDsbD (209). Interestingly, both studies propose four-state mechanisms, throughout which CcdA undergoes significant conformational changes to pass the reducing power from thioredoxin in the cytoplasm to thioredoxin-like proteins in the periplasm (208,209).

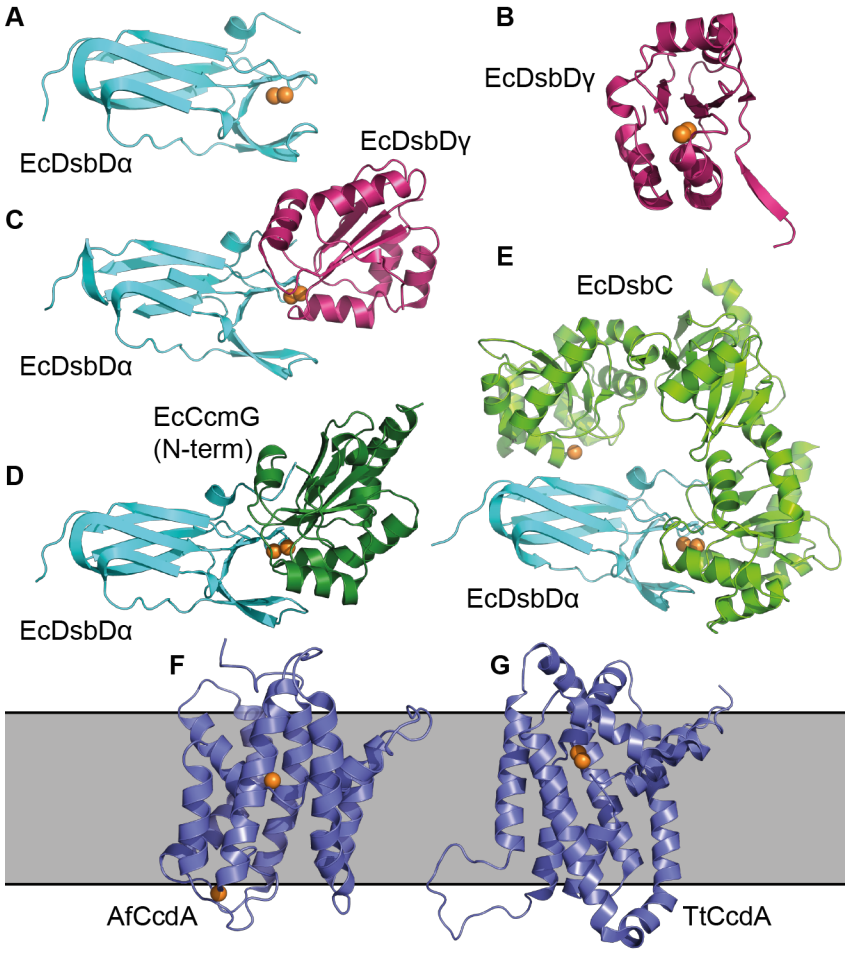


Figure 5.2: Structures of CcdA and the soluble domains of EcDsbD and their interactions. A. *Crystal structure of EcDsbDα (PDB ID: 1JPE).* ***B.*** *Crystal structure of EcDsbDγ (PDB: 1UC7).* ***C.*** *Crystal structure of the complex between EcDsbDα (cyan) and EcDsbDγ (dark pink) (PDB: 1VRS).* ***D.*** *Crystal structure of the complex between the N-terminal domain of EcCcmG (dark green) and EcDsbDα (cyan) (PDB: 1Z5Y).* ***E.*** *Crystal structure of the complex between EcDsbC (light green) and EcDsbDα (cyan) (PDB: 1JZD). NMR structures of CcdA from* ***F*** *A. fulgidus (PDB ID: 2N4X) and* ***G*** *T. thermophilus (PDB ID: 5VKV). In all structures, orange spheres represent the positions of the catalytic cysteines.*

5.1.2 PmScsB

Unlike CcdA, the predicted architecture of PmScsB is very similar to EcDsbD (210) with PmScsB predicted to have both periplasmic domains separated by 8 transmembrane helices (Figure 5.3). The major difference is the structure of the N-terminal or α- domain which, (as described in chapter 4), consists of two immunoglobulin-like folds in PmScsBα and only one in EcDsbDα. I used Phyre2 (211) to produce a homology model of the PmScsBγ domain and this model contains a thioredoxin fold like EcDsbDγ (63). Each domain of PmScsB also has a predicted redox active cysteine pair.

Given the predicted structural similarities between EcDsbD and PmScsB, we hypothesise that PmScsB will also transport electrons like EcDsbD. Functionally and structurally characterising PmScsB will provide more information about the diversity of membrane electron transporters and improve our understanding of the Scs proteins. The first step towards the characterisation of PmScsB is to develop an expression and purification protocol to produce high yields of the purified protein. This chapter describes the work performed towards achieving this goal.

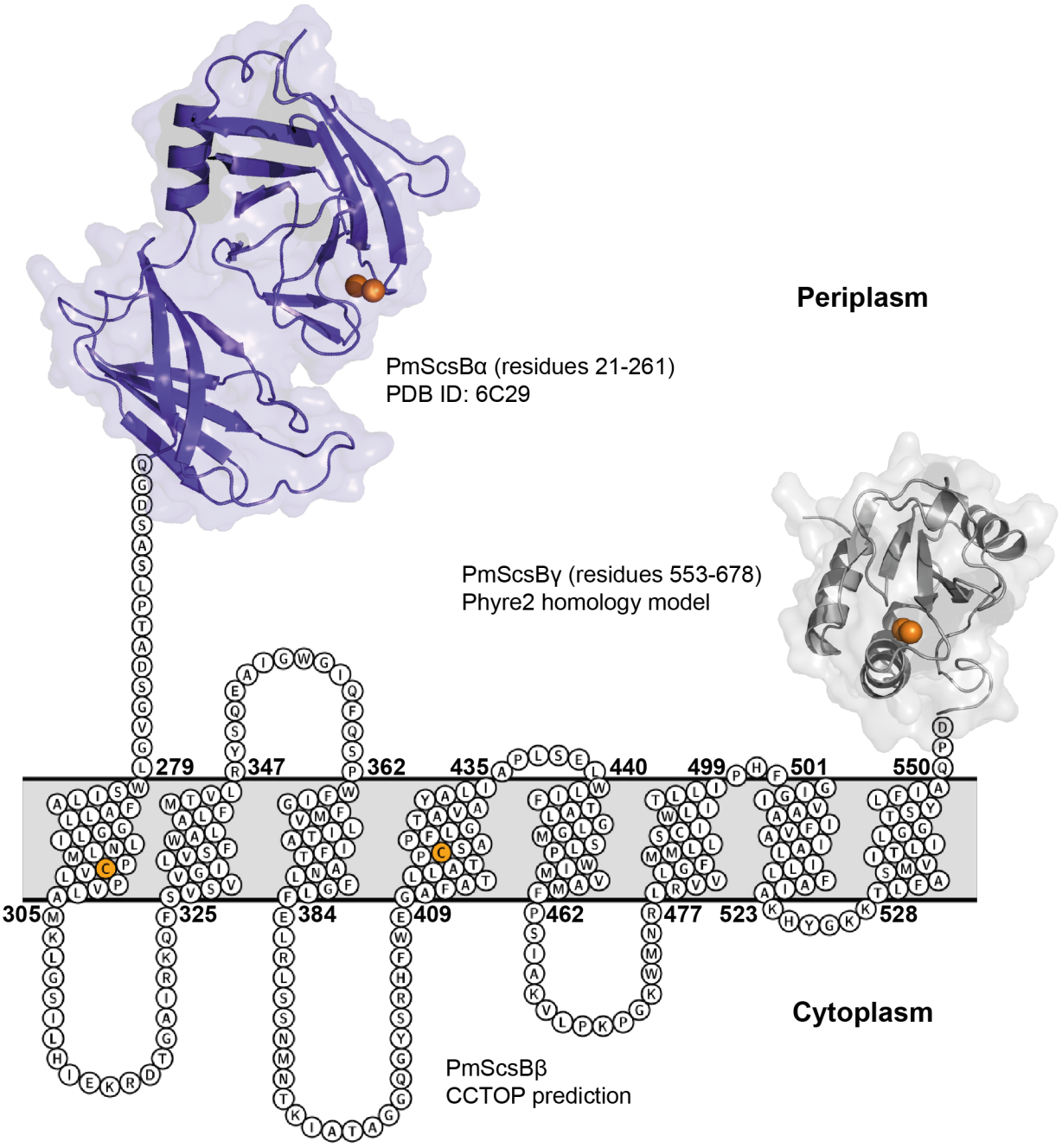


Figure 5.3: PmScsB. *The crystal structure of PmScsBα (purple, PDB ID: 6C29) has two immunoglobulin-like folds. Topology prediction using CCTOP* (210) *suggests PmScsBβ consists of eight transmembrane α-helices. The homology model of PmScsBγ, which I generated using Phyre2* (211)*, contains a thioredoxin fold. The topology diagram of PmScsBβ was made in Protter* (212)*.*

5.1.3 Membrane proteins

Working with membrane proteins presents many challenges. They are often poorly expressed, unstable and insoluble in aqueous solutions therefore requiring detergents to enhance their stability. Generally they are not as easily crystallised as soluble proteins and the limited number of membrane protein structures available in the Protein Data Bank (PDB) exemplifies this. As of the 16th of July 2018, only 3.4% of all PDB structures were membrane proteins (www.rcsb.org). Membrane proteins have very important biological roles however, and represent the largest class of drug targets (213,214). This highlights the need for more research on the structural and functional characterisation of these proteins. To enable this, specific methods and reagents have been developed to assist with membrane protein expression, purification and crystallisation.

One key development has been the membrane protein overexpression and purification pipeline developed by Drew et al. (215-217) that utilises Green Fluorescent Protein (GFP). By tagging membrane proteins with GFP, they can be more easily tracked throughout expression and purification, making screening and optimisation much more efficient (Figure 5.4). The combination of fluorescence detection and size-exclusion chromatography (FSEC) allows the monodispersity of the protein in detergent to be easily assessed prior to purification (218).

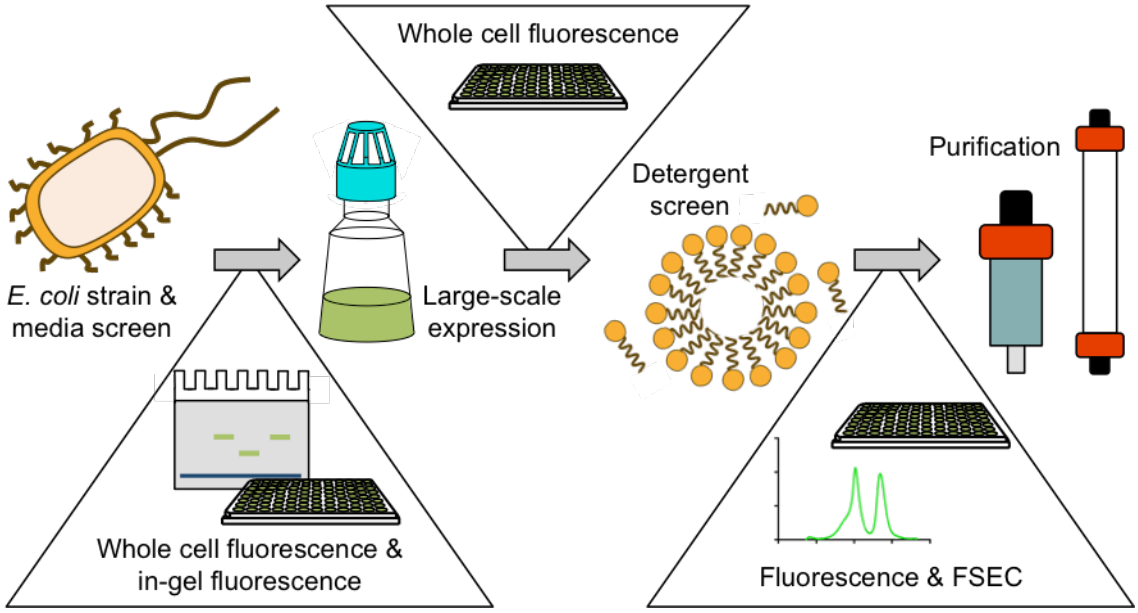


Figure 5.4: GFP-based membrane protein expression and purification pipeline. *Between the screening steps in the pipeline the fluorescence is measured to assess the quality and quantity of the membrane protein-GFP fusion.*

A limitation of the GFP-based pipeline is that the C-terminus of the membrane protein has to be cytoplasmic for GFP to fold (219). This is a problem for proteins like PmScsB that have a periplasmic C-terminus. However, a “superfolder” GFP (sfGFP) variant has been engineered, which contains multiple mutations that improve the folding of GFP (220) and has been shown to fold in the periplasm (221). Theoretically, sfGFP could be used as a tag on the periplasmic C-terminus of PmScsB.

Different strains of *E. coli* have also been developed specifically for membrane protein overexpression. Given that endogenous membrane proteins are typically at low abundance in cells, recombinant overexpression in *E. coli* makes the purification of high yields of membrane protein much easier. Unfortunately membrane protein overexpression in the classical T7-promoter expression strains (such as BL21 (DE3)) is often toxic to the cells and results in aggregation of the protein (222). This is thought to happen because the machinery that embeds the protein in the cytoplasmic membrane becomes saturated (222). The Walker strains, C41 (DE3) and C43 (DE3) are derivatives of the BL21 (DE3) *E. coli* strain, which have mutations that make them amenable to membrane protein overexpression (223). One of these mutations is in the *lac*UV5 promoter of the T7 RNA polymerase gene, which reduces the expression of the polymerase (224) but there are potentially other uncharacterised mutations in the strains. Another strain is Lemo21, which is the BL21 (DE3) strain containing pLemo (224). The plasmid pLemo harbours the gene encoding T7 lysozyme (inhibitor of T7 RNA polymerase) under an L-rhamnose inducible and titratable promoter (224,225). T7 lysozyme is also encoded on other plasmids like pLysS and pLysE (226), but pLemo is more favourable because it allows the optimisation of T7 lysozyme expression simply by varying the concentration of L-rhamnose in the media.

Detergents are used throughout membrane protein purification to solubilise and keep membrane proteins stable in solution. Detergents generally have a hydrophilic head group and a hydrophobic tail but there is a broad range of commercially available detergents and these vary in structure and chemical properties (Figure 5.5). Some detergents, such as *n*-dodecyl-β-D-maltopyranoside (DDM), are considered to be mild because of their large, non-ionic head groups. Other detergents, such as *N,N*-dimethyldodecylamine *N*-oxide (LDAO), are harsher because of their smaller, ionic head groups (227). Detergents can exist in solution as monomers and as micelles, though micelle formation only occurs above the critical micelle concentration (CMC). The micelle size can impact on crystal formation,

as larger micelles prevent the protein from making crystal contacts, therefore, new detergents have been designed to increase the stability of solubilised membrane proteins and overcome problems associated with crystallisation (227). Detergent screening techniques, such as the unfolding assay utilising the dye N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]-maleimide (CPM) (228), have also been developed to identify optimal detergents for stabilisation of different membrane proteins (229).

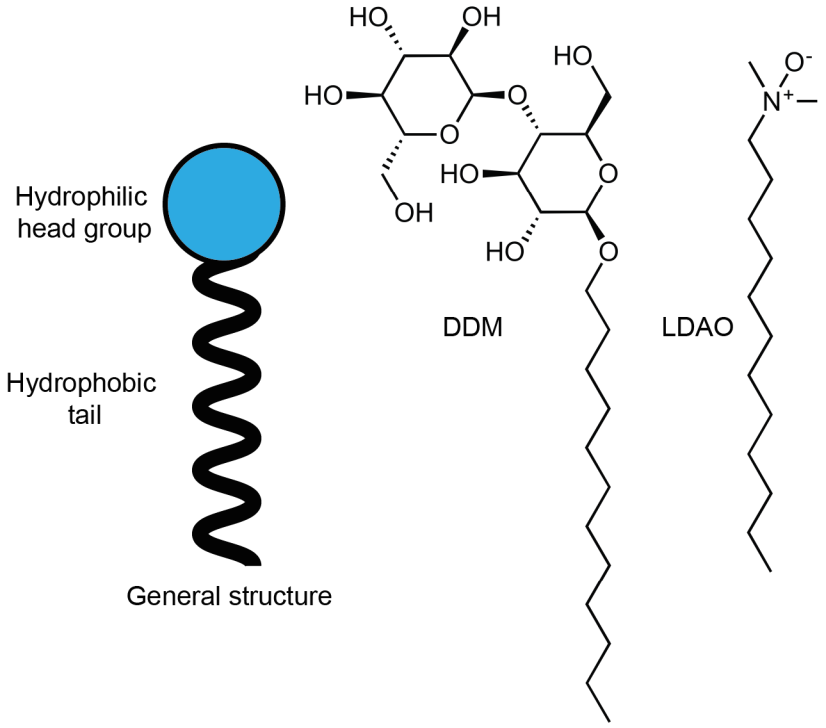


Figure 5.5: Structure of detergents. *Detergents generally consist of a hydrophilic head group and a hydrophobic tail, however the chemical properties of these components can vary significantly. DDM and LDAO are shown in the figure to exemplify the diversity in detergent head group structure.*

Detergents, however, do not replicate the membrane environment very well, so some membrane proteins may not be stable in detergent micelles. Membrane mimetics, such as liposomes and nanodiscs can instead be used in the characterisation of membrane proteins (230,231). Lipidic cubic phase (LCP), which is a membrane-mimetic matrix, can also be used for crystallisation in combination with specific membrane protein crystallisation screens that have been developed (232). The advances in single particle analysis with cryo-electron microscopy have also assisted in membrane protein structure determination, as this technique does not require crystals or large amounts of material, which are needed for crystallography (233,234).

This chapter details the steps I have taken toward the expression and purification of the full-length transmembrane protein, PmScsB. A construct encoding PmScsB with a C-terminal sfGFP tag and His_8-tag was designed and a GFP-based approach (217) was used to screen different expression strains, media and suitable detergents for solubilisation. Large-scale expression was performed and purified PmScsB was obtained through the use of immobilised metal affinity chromatography and size-exclusion chromatography. These findings will assist in future work on the characterisation of this integral membrane protein.

5.2 Methods

5.2.1 Generation of the PmScsB-sfGFP expression construct

The PmScsB gene (encoding Uniprot ID: B4EV20) was codon optimised for expression in *E. coli* using GeneOptimizer (Thermo Fisher Scientific, Massachusetts, USA). The gene without the predicted periplasmic signalling sequence (residues 1-21) was ordered from GENEWIZ (New Jersey, USA) and inserted into pET20b(+) (Novagen®, Merck, Darmstadt, Germany) which contains a PelB leader sequence allowing targeted expression of the protein to the periplasm of *E. coli* (235). Using this construct, it was difficult to determine expression levels of PmScsB and trace the protein throughout the purification process. After personal communication with Dr David Drew (Stockholm University), a new strategy was employed and primers were designed (Table 5.1) to incorporate the native predicted periplasmic signalling sequence of PmScsB and to insert the gene into pWaldo-sfGFPd (plasmid provided by Dr David Drew, Stockholm University, Sweden) (217,220,236). The plasmid pWaldo-sfGFPd is a protein expression construct that contains superfolder GFP (sfGFP) downstream of restriction enzyme sites (Figure 5.6A). Overlap extension PCR was used to mutate an NdeI cleavage site in the PmScsB gene and the gene was then amplified with primers encoding the signal peptide sequence and NdeI and KpnI cleavage sites. The PmScsB PCR product and pWaldo-sfGFPd vector were cut with the restriction enzymes and ligated according to the manufacturers instructions (New England Biolabs, Massachusetts, USA). The ligation product was transformed into *E. coli* TOP10 (Invitrogen, California, USA) competent cells and plated on lysogeny broth (LB) agar supplemented with 50 µg/ml kanamycin. Resulting colonies were cultured in LB with 50 µg/mL kanamycin and the construct was extracted using the QIAGEN (Hilden, Germany) miniprep kit. The final construct, pEF014, contained the codon optimised PmScsB gene (encoding Uniprot ID: B4EV20, residues 1-678) followed by a TEV protease cleavage site, superfolder GFP and a His_8 tag and this was confirmed by sequencing (Figure 5.6B). After TEV protease cleavage there are an additional 11 amino acids (RVPGSENLYFQ) on the C-terminal end of the protein.

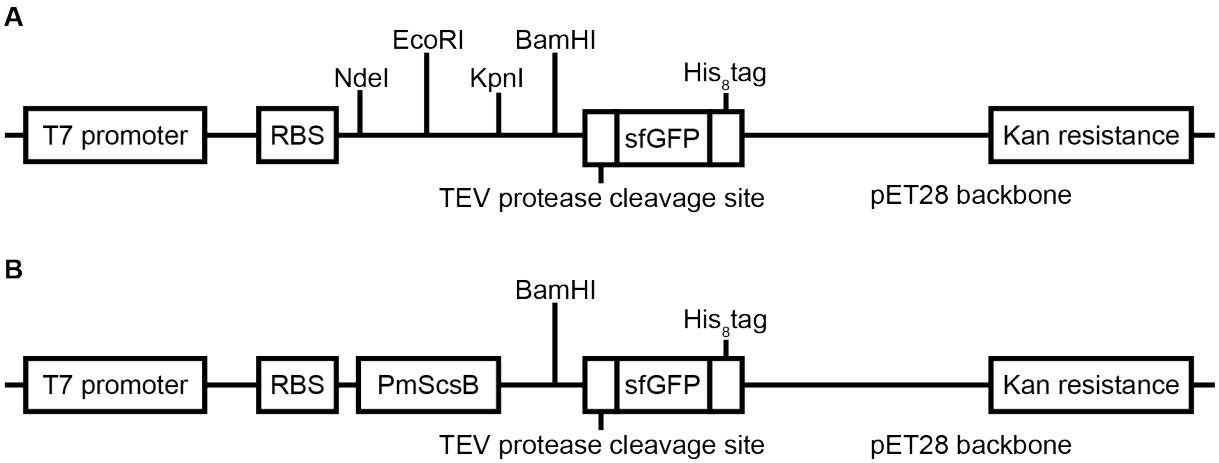


Figure 5.6: Visual representation of the constructs used in chapter 5. A. *pWaldo-sfGFPd construct contains sfGFP downstream from restriction enzyme sites.* ***B.*** *The pEF014 construct was made for the expression of PmScsB-sfGFP. The elements on the constructs are not drawn to scale and the schematic was adapted from diagrams published previously* (217)*.*

Table 5.1: Primers used in chapter 5. *Mutated and overhanging nucleotides are shown in bold.*

Name	Description	Sequence (5'➔3')
EF-028	Rev NdeI cut site mutation	GCGATCAGGGCATA**G**GCAACTGCGGTGCC
EF-029	Fwd NdeI cut site mutation	GGCACCGCAGTTGC**C**TATGCCCTGATCGC
EF-030	Rev PmScsB with KpnI cut site	**GCGGCGGGTACCCG**CTGTGCGCCTTTTG
EF-031	Fwd PmScsB signal sequence (1)	**CTGATTGCACTGATGCTGATGCTGTTTAGCCTGAGCAG CCTGGCA**GCAGATACCGGTTGG
EF-032	Fwd PmScsB signal sequence (2) with NdeI site	**CGCCGCCATATGGTTCGCAGCTTT**CTGATTGCACTGAT GC

5.2.2 Small-scale expression screen

Test expression of the PmScsB-sfGFP construct in various *E. coli* strains and media was carried out to determine the optimal strain and media for large-scale expression. Three *E. coli* strains, (LEMO21, BL21 (DE3) pLysS and C43, Table 5.2) and three types of media

(LB and two types of autoinduction media, ZYP and PASM, Table 5.3) were screened. Fresh transformations of pEF014 in the different strains of *E. coli* were used to inoculate 10 mL aliquots of LB, with 50 µg/mL kanamycin and, if appropriate, 34 µg/mL chloramphenicol. Cultures were grown overnight at 37°C. For the test expressions, 200 µL of the overnight cultures was added to 10 mL (1 in 50 dilution) of each type of media, containing the appropriate antibiotics. The test cultures were incubated at 37°C until an OD_{600} of ~0.4-0.5 (1-2 h) was reached. At this point, the incubation temperature was lowered to 25°C and the LB cultures were induced with 0.4 mM IPTG. All cultures were grown overnight. After approximately 22 h post induction, the OD_{600} was measured and 1 mL samples of each condition were taken. The cells in these samples were spun down, washed in 1×PBS and resuspended in 100 µL of 1×PBS. These 100 µL suspensions were added to Greiner black 96-well plates with clear flat bottom wells (Merck, Darmstadt, Germany). The whole cell fluorescence was measured in a Synergy H1 Hybrid plate reader (BioTek, Vermont, USA) using the bottom read mode, a gain of 60 and excitation and emission wavelengths of 485 and 520 nm, respectively. The fluorescence reported for each small-scale sample was adjusted for the background fluorescence of PBS and normalised to an OD_{600} of 1. The whole cells were also run on a Novex 4-12% Tris-Glycine SDS-PAGE gel (Thermo Fisher Scientific, Massachusetts, USA) and analysed using a Bio-Rad (California, USA) ChemiDoc™ MP imaging system with a 530/28 nm filter and blue epi-illumination setting.

Table 5.2: **E. coli *strains screened.***

Strain	Description	Source
Lemo21 (DE3)	BL21 (DE3) strain with pLemo, which is maintained by chloramphenicol	New England Biolabs (Massachusetts, USA)
BL21 (DE3) pLysS	BL21 (DE3) strain containing the plasmid pLysS, which is maintained by chloramphenicol	Invitrogen Australia
C43 (DE3)	Strain derived from BL21 (DE3) with mutations that make it more capable of membrane protein overexpression	Lucigen (Wisconsin, USA)

Table 5.3: Types of media used in expression screening. *ZYP and PASM were prepared based on the protocols described previously (139). For maintenance of pEF014, the media was supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol was also added to the media for growth of Lemo21 and BL21 (DE3) pLysS strains.*

Media	Components	Induction
LB	• 0.5% w/v yeast extract • 1% w/v tryptone • 0.5% w/v NaCl	IPTG
ZYP	• ZY media (0.5% w/v yeast extract, 1% w/v tryptone) • 1× 5052 (0.5% w/v glycerol, 0.05% w/v glucose, 0.2% w/v α-lactose) • 1× NPS (25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4) • 1× Trace metals mix (50 µM $FeCl_3•6H_2O$, 20 µM $CaCl_2$, 10 µM $MnCl_2•4H_2O$, 10 µM $ZnSO_4•7H_2O$, 2 µM $CoCl_2•6H_2O$, 2 µM $CuCl_2•2H_2O$, 2 µM $NiCl_2•6H_2O$, 2 µM $Na_2MoO_4•5H_2O$, 2 µM $Na_2SeO_3•5H_2O$, 2 µM H_3BO_3) • 1 mM $MgSO_4$	Autoinduction
PASM	• 1× 5052 (0.5% w/v glycerol, 0.05% w/v glucose, 0.2% w/v α-lactose) • 1× NPS (25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4) • 0.2× Trace metals mix (10 µM $FeCl_3•6H_2O$, 4 µM $CaCl_2$, 2 µM $MnCl_2•4H_2O$, 2 µM $ZnSO_4•7H_2O$, 0.4 µM $CoCl_2•6H_2O$, 0.4 µM $CuCl_2•2H_2O$, 0.4 µM $NiCl_2•6H_2O$, 0.4 µM $Na_2MoO_4•5H_2O$, 0.4 µM $Na_2SeO_3•5H_2O$, 0.4 µM H_3BO_3) • 1 mM $MgSO_4$ • 200 µg/mL of each amino acid (except cysteine) • 0.1 µM Vitamin B12	Autoinduction

5.2.3 Large-scale expression

Two conditions were selected from the small-scale expression screen to scale up. These were *E. coli* Lemo21 in ZYP and *E. coli* C43 in LB, with 0.4 mM IPTG induction. Fresh transformants were used to grow 100 mL overnight cultures in LB, supplemented with

antibiotics. 20 mL of these overnight cultures was used to inoculate 2.5 L baffled shaker flasks containing 1 L of the appropriate expression media with antibiotics. These were incubated at 37°C, 220 rpm to an OD_{600} of ~0.4 (1-2 h), before the temperature was reduced to 25°C and the LB cultures were induced with 0.4 mM IPTG. The expression cultures were grown overnight before the cells were harvested by centrifugation and stored at -80°C. OD_{600} and whole cell fluorescence was measured to determine expression levels of PmScsB-sfGFP.

5.2.4 Membrane preparation

The harvested cells were resuspended in 3 mL of 1× PBS per gram of cell pellet and protease inhibitor (1 in 1000 dilution in lysate, BioPioneer, Inc., USA) and DNase (3.3 µg/ml final concentration) were added. The cells were lysed using a Constant Systems Cell Disruptor with two passes at 25 kpsi. The lysate was centrifuged at 10,000 rpm, 15 min, 4°C to remove the unlysed cells and cell debris. Ultracentrifugation at 42,500 rpm, 1 h, 4°C was used to separate the membranes from the soluble fraction. The membranes were resuspended in 1× PBS to a final volume of 15 mL per litre of initial expression culture, flash frozen in liquid nitrogen and stored at -80°C. The fluorescence of 100 µL of the 1× PBS membrane suspension was measured as described for the whole cell measurement.

5.2.5 Detergent screen

Eight different detergents (Table 5.4) were screened for their ability to solubilise PmScsB-sfGFP. The membranes were diluted to 20,000 RFU and 100 µL of 10% w/v or v/v detergent was added to 900 uL of membrane suspension to give a final detergent concentration of 1%. The membrane-detergent mixtures were incubated at 4°C for 2 h, mixing on a roller mixer. The unsolubilised membranes were removed by ultracentrifugation (29,300 rpm, 4°C for 30 min) and the fluorescence of 100 µL of the detergent solubilised material was analysed as previously described. The three detergent solubilised samples that had the highest fluorescence reading were then analysed by FSEC to determine if PmScsB-sfGFP was stable. One hundred µL from each of the crude LDAO-, UDM- and DDM-solubilised membranes was applied to a Superose 6 Increase 10/300 column (GE Healthcare, Chicago, USA) on an ÄKTApurifier (GE Healthcare, Chicago, USA) with an in-line Jasco (Maryland, USA) FP-2020 plus fluorescence detector (Excitation: 488 nm, Emission: 512 nm). The buffer used for the FSEC analysis was 1× PBS with the addition of the detergent of interest at a concentration of 5 times the CMC.

Table 5.4: Detergents screened for solubilisation efficiency of PmScsB-sfGFP.

Abbreviation	Full name
DM	*n*-Decyl-β-D-maltopyranoside
DDM	*n*-Dodecyl-β-D-maltopyranoside
NM	*n*-Nonyl-β-D-maltopyranoside
UDM	*n*-Undecyl-β-D-maltopyranoside
OG	*n*-Octyl-β-D-glucopyranoside
LDAO	*N,N*-Dimethyldodecylamine *N*-oxide
LMNG	2,2-Didecylpropane-1,3-bis-β-D-maltopyranoside / Lauryl Maltose Neopentyl Glycol
Triton-X	Polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether

5.2.6 Purification

The membranes (prepared as described previously) at a fluorescence of ~20,000 RFU were solubilised using 1% w/v DDM in 1× PBS and 20 mM imidazole. The solubilisation mixture was incubated at 4°C, stirring gently, for 2 h after which ultracentrifugation at 29,300 rpm, 4°C for 30 min was used to pellet any unsolubilised membranes. Using a peristaltic pump, the supernatant was applied, at approximately 1 mL/min, over a 5 mL His-trap (GE Healthcare, Chicago, USA) nickel column equilibrated with 1× PBS, 20 mM imidazole to capture the His-tagged PmScsB-sfGFP. The column was then washed with 5 column volumes each of IMAC wash buffers 1 and 2 (Table 5.5) and the captured protein was eluted from the column in 10 mL of IMAC elution buffer. Tobacco etch virus (TEV) protease was added to the eluted protein, in an approximately 1:1 (TEV protease: protein) molar ratio to remove the sfGFP-His_8tag from PmScsB. The mixture was dialysed in 3kDa cut-off dialysis tubing suspended in 2 L of dialysis buffer (Table 5.5) by stirring overnight at 4°C. The protein was removed from the dialysis tubing and imidazole concentration adjusted to 20 mM. The 5 mL His-trap nickel column was washed, equilibrated and reverse IMAC was performed to remove any His-tagged contaminants. The flow through from the reverse IMAC, which should contain PmScsB, was then concentrated to 5 mL using a 100 kDa cut-off centrifugal concentrator (Merck, Darmstadt, Germany) and injected into an FPLC for purification via SEC. The protein was run over a Superdex 200 16/60, with SEC buffer. Peak fractions were collected and concentrated to ~5 mg/mL using 100 kDa cut off centrifugal concentrators. Throughout the purification, protein concentrations were determined by measuring the A_{280} using a Nanodrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific, Massachusetts, USA) and

adjusting for the predicted extinction coefficient (Table 5.6) (140). Samples from each stage of the purification were analysed with SDS-PAGE, run on a 4-12% Tris-Glycine gel and stained with Coomassie blue. This was also used to assess the purity of PmScsB. The fluorescence of 100 µL samples taken at appropriate stages of the purification was also analysed as described previously.

Table 5.5: Buffers used in PmScsB purification.

Buffer	Components
1× PBS	2.7 mM KCl, 137 mM NaCl, 1.76 K_2PO_4 pH 7.4
IMAC wash buffer 1	20 mM Imidazole, 1× PBS, 0.05% w/v DDM
IMAC wash buffer 2	40 mM Imidazole, 1× PBS, 0.05% w/v DDM
IMAC elution buffer	250 mM Imidazole, 1× PBS, 0.05% w/v DDM
Dialysis buffer	1× PBS, 0.03% w/v DDM
SEC buffer	20 mM HEPES pH 7.5, 150 mM NaCl, 0.01% w/v DDM

Table 5.6: Theoretical molecular weight and the absorbance at 280 nm of a 1 mg/mL solution (Abs 0.1%) of different PmScsB-sfGFP products (140).

Protein	Molecular Weight (kDa)	Abs 0.1% (reduced/oxidised)
*PmScsB-sfGFP-His	101.4	1.561/1.567
*PmScsB-	73.2	1.903/1.909
-sfGFP-His	28.1	Not necessary

*With no signalling sequence (residues 1-20), as this should be cleaved.

5.2.7 Mass spectrometry analysis of the purified protein

A 5 mg/ml sample of the purified protein was sent for mass spectrometry analysis to confirm its identity. Dr Gordon King (UQROCX) performed this work by digesting the protein with trypsin and analysing the molecular weight of the peptides using matrix assisted laser desorption/ionisation (MALDI). Tandem mass spectrometry (MS/MS) was used to determine the sequence of some peptides. The observed masses and sequences were compared to those peptides predicted by the *in silico* digest of PmScsB with trypsin (PeptideMass tool, ExPASy (140)).

5.3 Results

5.3.1 Small-scale expression screen

Three different *E. coli* strains and three different types of media were screened, to find suitable conditions for the expression of ScsB. The three strains screened were, Lemo21 (DE3), C43 (DE3) and BL21 (DE3) pLysS and the types of media screened were, LB with 0.4 mM IPTG induction and ZYP and PASM, a rich and minimal autoinduction media, respectively. Out of the nine different combinations, two conditions, Lemo21 (DE3) grown in ZYP and C43 (DE3) grown in LB with 0.4 mM IPTG induction, had a much higher level of whole cell fluorescence than the others (Figure 5.7A). To ensure the fluorescence was not a consequence of free GFP, samples were taken for analysis by SDS-PAGE in-gel fluorescence. The results from this analysis suggested that most of the conditions, including the top two, contained a fluorescent protein at the approximate size of PmScsB-sfGFP (molecular weight = 101.4 kDa) as well as some free GFP (molecular weight = 28.1 kDa) (Lane 4 and 8, Figure 5.7B).

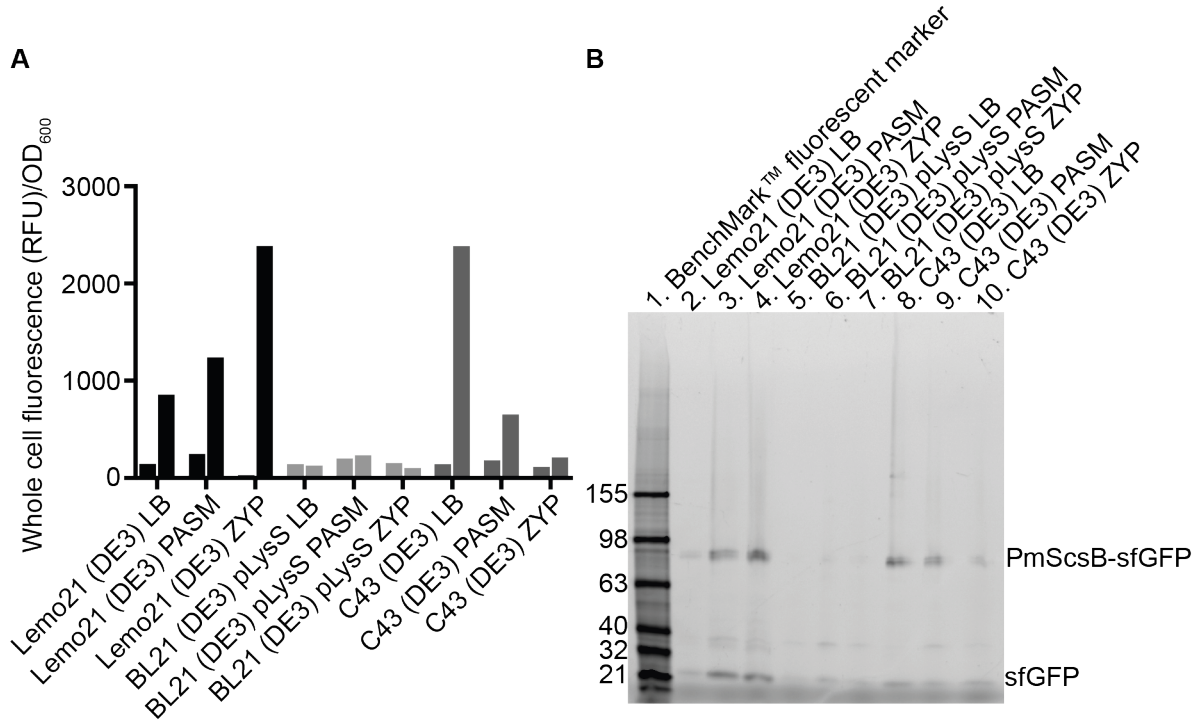


Figure 5.7: Results from the small-scale expression screen. A. *Whole cell fluorescence for the conditions screened, normalised to an OD_{600} of 1. The two bars for each condition represent the fluorescence at time 0 (i.e. at induction) and after 22 hours, respectively.* ***B.*** *In-gel fluorescence of samples taken from each of the screening conditions. The expected molecular weight of PmScsB-GFP-His is 101.4 kDa, however, membrane proteins typically run at a lower molecular weight than expected as they are not fully denatured (heated) for SDS-PAGE.*

5.3.2 Expression scale up

The two conditions identified in 5.3.1 were tested at a 1 L scale to determine if the conditions were amenable to scale up. PmScsB expression using the C43 (DE3) strain in LB with 0.4 mM IPTG induction resulted in a final OD_{600} of ~4.8 and a fluorescence reading of ~2,300 RFU. The expression in Lemo21 using ZYP gave a final OD_{600} of ~11.2 and a fluorescence reading of ~22,000 RFU. The Lemo21 cells were therefore chosen for membrane preparation and subsequent large-scale expression of PmScsB used the combination of Lemo21 with ZYP media.

5.3.3 Detergent screening

Membrane proteins require detergents for their solubilisation and subsequent purification. There are many detergents that are commercially available and can be assessed to determine the optimal detergent that extracts a significant amount of the protein from the membrane whilst maintaining the protein's stability. In the detergent screen for PmScsB, eight different detergents were assessed for their ability to extract PmScsB-sfGFP from the membrane preparation by measuring the fluorescence of the solubilised material. Initial tests showed that LDAO had the highest fluorescence, followed by UDM and DDM (Figure 5.8A), so the solubilised material produced by these three detergents was further analysed by FSEC. The large peak in the UV trace at ~10 mL in the LDAO chromatogram suggested that this detergent produced aggregated material (Figure 5.8B). The fluorescence trace also had a large peak corresponding to free GFP (elution volume ~19 mL) and only a small peak around 15 mL, which is approximately the expected elution volume for PmScsB-sfGFP in a detergent micelle (Figure 5.8B). The chromatogram for the UDM solubilised membranes also suggested that aggregation and free GFP were produced, but the peak at ~15 mL was much larger (Figure 5.8C). DDM produced the best FSEC profile, out of those detergents analysed, with a larger, almost monodisperse fluorescent peak at ~15 mL, which also correlated to a large peak in the UV trace (Figure 5.8D). Therefore, for further work DDM was used to solubilise PmScsB-sfGFP and was maintained at concentrations above the CMC in the purification buffers to stabilise the protein.

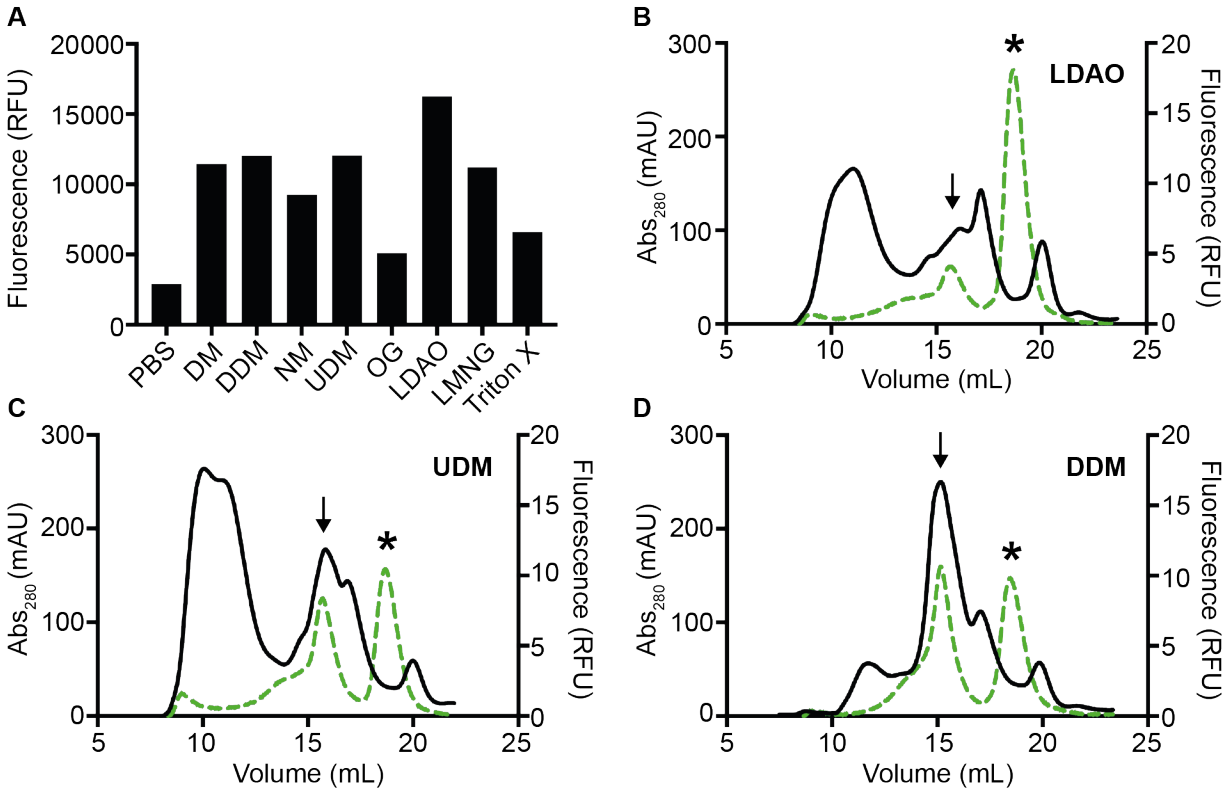


Figure 5.8: PmScsB detergent screen. A. *Fluorescence of solubilised membranes using different detergents.* ***B-D.*** *FSEC chromatograms of the crude solubilised membranes from three of the screened detergents. The solid black line is the UV trace and the dashed green line shows the detected fluorescence. For each FSEC chromatogram, an arrow and an asterisk indicate the peaks that are predicted to correspond to PmScsB-sfGFP and sfGFP, respectively.*

5.3.4 Purification

PmScsB was successfully purified from expression in *E. coli* Lemo21 (DE3) using four steps; (1) IMAC to separate PmScsB-sfGFP from other solubilised membrane proteins, (2) TEV cleavage to remove the sfGFP and His_8tag and (3) reverse IMAC to remove His-tagged and other Nickel binding contaminants followed by (4) SEC. On the SDS-PAGE gel, no band corresponding to PmScsB-sfGFP can be seen in solubilised membranes or the IMAC flow through and washes (Figure 5.9). The IMAC step is effective at extracting the fusion protein as evidenced by (i) the strong band corresponding to the molecular weight of PmScsB-sfGFP in the elution sample on SDS-PAGE (Lane 6, Figure 5.9A), (ii) the green colour of the eluted protein (Figure 5.9B) and (iii) the decrease in RFU units from ~15,000 RFU in the solubilised membranes to ~4,700 RFU in the IMAC flow through. TEV protease cleavage was also successful, as the band corresponding to PmScsB-sfGFP is not present in the post-TEV cleavage sample, but a slightly smaller (less ~28 kDa) band is

present (Lane 7, Figure 5.9). Reverse IMAC removed most of the contaminants, though some PmScsB was present in the reverse IMAC elution and after injection onto SEC, step 4, the chromatogram obtained shows one broad peak with a tail (Figure 5.10). The SEC purification step appears to have resulted in a large loss of protein as close to 10 mg of protein was injected but only ~500 µg was obtained from the pooled peak fractions (purification from 2.5 L of expression culture, ~ 40 g of *E. coli* pellet). Overall yield of the purification was ~ 200 µg per L of expression culture.

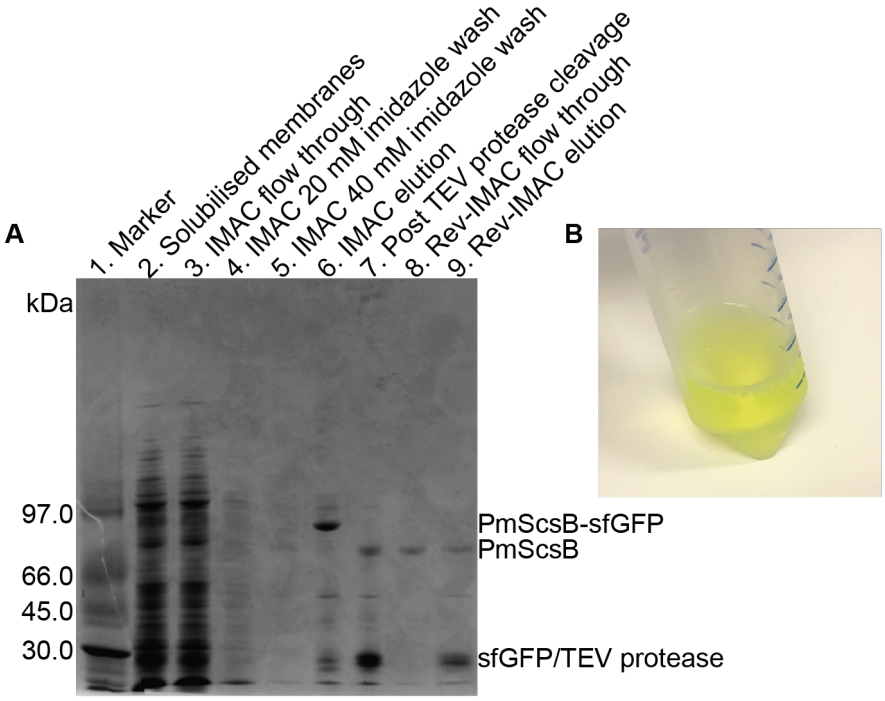


Figure 5.9: Purification of PmScsB. A. *SDS-PAGE analysis of samples taken throughout the purification of PmScsB.* ***B.*** *Pooled eluted protein from the IMAC purification step (Lane 6).*

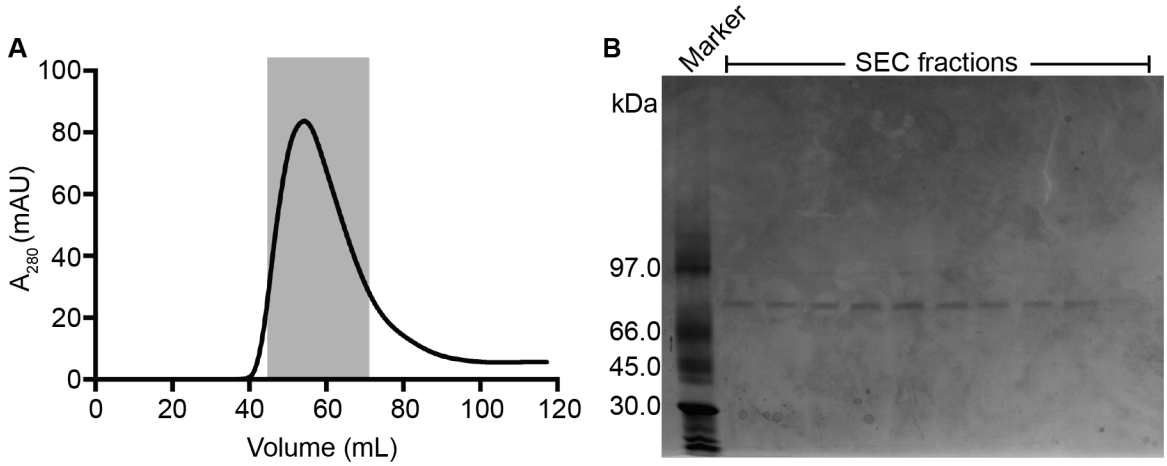


Figure 5.10: SEC step in the purification of PmScsB. A. *Chromatogram of the SEC purification step of PmScsB* ***B.*** *SDS-PAGE analysis of the fractions in the shaded region on the SEC chromatogram (**A**).*

5.3.5 Mass spectrometry analysis of the purified protein

Mass spectrometry analysis, involving a trypsin digest in combination with MALDI, was used to determine if the purified protein contained PmScsB. Six of the peptides observed in the MALDI spectrum had masses correlating to those predicted by an *in silico* trypsin digest of PmScsB (Figure 5.11). The full sequence of two of these peptides was confirmed by MS/MS (Table 5.7). The observed masses also correlated to predicted peptides in all domains of the PmScsB protein (Table 5.7).

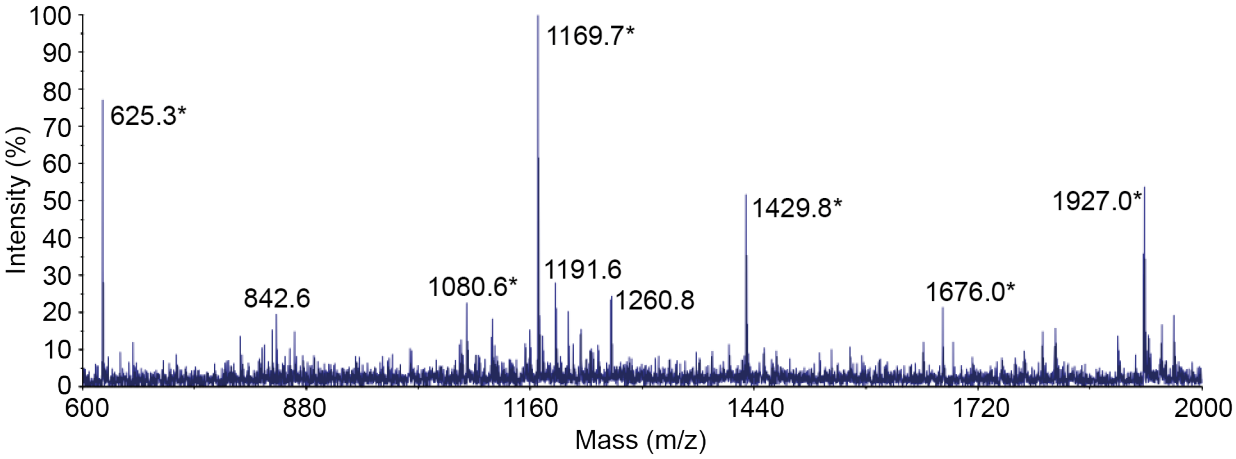


Figure 5.11: MALDI spectrum. *Spectrum obtained from a MALDI experiment on the trypsin-digested purified protein. Those masses with an asterisk (*) agree with predicted peptides from a trypsin digest of PmScsB.*

Table 5.7: Correlation between the observed and predicted masses of peptides generated by a trypsin digest of PmScsB. *The sequences shown in bold were confirmed by MS/MS.*

Observed mass (m/z)	**Predicted mass (monoisotopic) (m/z) (140)**	**Sequence and position in PmScsB**	**Missed cleavages?**
625.3	625.3	62-TYWR-65	No
1080.6	1080.5	395-IATAGGQGYSR-405	No
1169.7	1169.7	610-**ALSEPDVVALR**-620	No
1429.8	1429.7	244-**ALSFVVTDDGYSR**-256	No
1676.0	1675.9	596-ANKYNVLLRDDVQK-609	2
1927.0	1926.9	21-ADTGWLTM**PDNDHAQVR**-37	No

5.4 Discussion

The expression and purification of PmScsB is the first step in structurally and functionally characterising this protein. The results from this chapter indicate that preparation of the transmembrane protein PmScsB can be achieved through expression of the PmScsB-sfGFP construct in Lemo21 (DE3) using ZYP media, solubilisation of the membranes in DDM and purification using IMAC and SEC. Mass spectrometry analysis confirmed that the purified protein is PmScsB. These data suggested that the predicted periplasmic signalling peptide (residues 1-20) of PmScsB is removed during expression. The mass of 1927.0 correlates to residues 21-37 of PmScsB, however there is no trypsin cleavage site between residues 20 and 21, which suggests that this site must be cleaved by the signal peptidase during translocation.

The fusion of PmScsB with sfGFP allowed the expression of PmScsB to be readily optimised. The small-scale screen of different expression strains and media allowed the determination of two possible expression conditions, which was reduced to one after scale-up of the expression to 1 L. The strain and media combination that was chosen for further work was Lemo21 (DE3) in ZYP autoinduction media. Lemo21 (DE3) is a BL21 strain with the plasmid pLemo, which harbours the T7 Lysozyme gene, under an inducible and titratable L-rhamnose promoter. In the screening performed, no L-rhamnose was added to the media meaning that little, if any T7 Lysozyme would have been expressed. Screening different concentrations of L-rhamnose in the expression media may further optimise the expression of PmScsB. The temperature and length of expression are two other parameters that were not screened and could be varied to increase the expression of PmScsB.

The detergent solubilisation screen identified DDM as an appropriate detergent for PmScsB solubilisation. LDAO resulted in the highest solubilised membrane fluorescence but is generally considered quite harsh because of its ionic head group and smaller micelles (227) and unsurprisingly, PmScsB was not stable after extraction in this detergent. UDM gave a reasonable FSEC profile but DDM was chosen for further work, as it resulted in less aggregation seen on the UV trace. Unfortunately the large micelle size of DDM is considered to be detrimental for crystallisation (237,238), so further work could follow up on the other detergents, which were not analysed by FSEC. Of course only 8 detergents were screened in this study and many others are available that may produce better results. It is also possible that different detergents may be better for different stages

of the purification protocol and the stability of PmScsB in different detergents should be analysed using the unfolding assay with CPM dye. The detergent to protein ratio in the solubilisation step was not investigated in this preliminary work, but could impact on the amount of PmScsB solubilised. Alternative methods of membrane protein stabilisation, such as amphipols (239) and nanodiscs (230), are also worth investigation.

During membrane protein purification, the folding of the protein and GFP can sometimes hinder interaction of the His-tag with resin but in this case an IMAC purification step efficiently separated PmScsB-sfGFP from the rest of the solubilised membranes. Modifications of the IMAC buffers and overall procedure may be possible to improve the yield obtained from IMAC. Allowing greater residence time with the binding to the nickel column or recycling flowthrough back over the column could potentially improve the yield from this step. Changing to a batch method could also be investigated as another method of efficiently extracting PmScsB-sfGFP from the crude solubilised membranes. There is also the possibility of modifying the IMAC resin and buffer conditions to achieve increased binding, stability and ultimately protein yield.

TEV protease cleavage is often a problem in membrane protein purification because the detergent interferes or the cleavage site is inaccessible, however, this was not the case for the purification of PmScsB. TEV protease cleavage, overnight at 4°C, worked well, with very little uncleaved protein observed in the TEV protease treated sample on SDS-PAGE. The reverse IMAC removed many of the contaminants though a band corresponding to PmScsB was still observed in the reverse IMAC elution, suggesting that more washing or altering the buffer conditions may be beneficial to increase yields.

The SEC chromatogram only had one peak and analysis by SDS-PAGE revealed that the major component corresponded to PmScsB. This peak was very broad, which may be (i) a result of PmScsB being conformationally heterogeneous given the transporter action of the membrane domain and two potentially mobile periplasmic domains or (ii) indicate that the DDM micelles are of varying sizes. The SEC peak was not very intense and a significant proportion of protein was lost during this purification step. Though no aggregation was seen on the chromatogram, the concentration of DDM was lowered in the SEC running buffer to reduce the formation of detergent only micelles during the final concentration step. This could have reduced the stability of the protein and it may also be possible that 20 mM HEPES pH 7.5 is not the optimal buffer for purification of PmScsB. Analysis of

PmScsB stability in various buffers would assist in determining suitable conditions for the purification of this protein.

As described above there are many ways in which the expression and purification of PmScsB could be further optimised but this initial purification strategy is promising. The current protocol could be scaled up to produce enough protein for crystallisation screening and the yield is sufficient for analysis by electron microscopy. Future work could also involve assays to assess the activity of the purified protein and experiments to determine the stoichiometry of the full-length PmScsB-PmScsC complex.

Chapter 6

6. Conclusions and Future Directions

The suppressor of copper sensitivity (Scs) proteins are Dsb-like proteins involved in the bacterial response to copper stress. Very little is known about the Scs proteins and prior to this work these proteins had only been studied in two organisms: *S.* Typhimurium and *C. crescentus*. The work outlined in this thesis focused on the characterisation of two of the four proteins - ScsC and ScsB - from the clinically important human uropathogen *P. mirabilis*. This work is summarised in Figure 6.1.

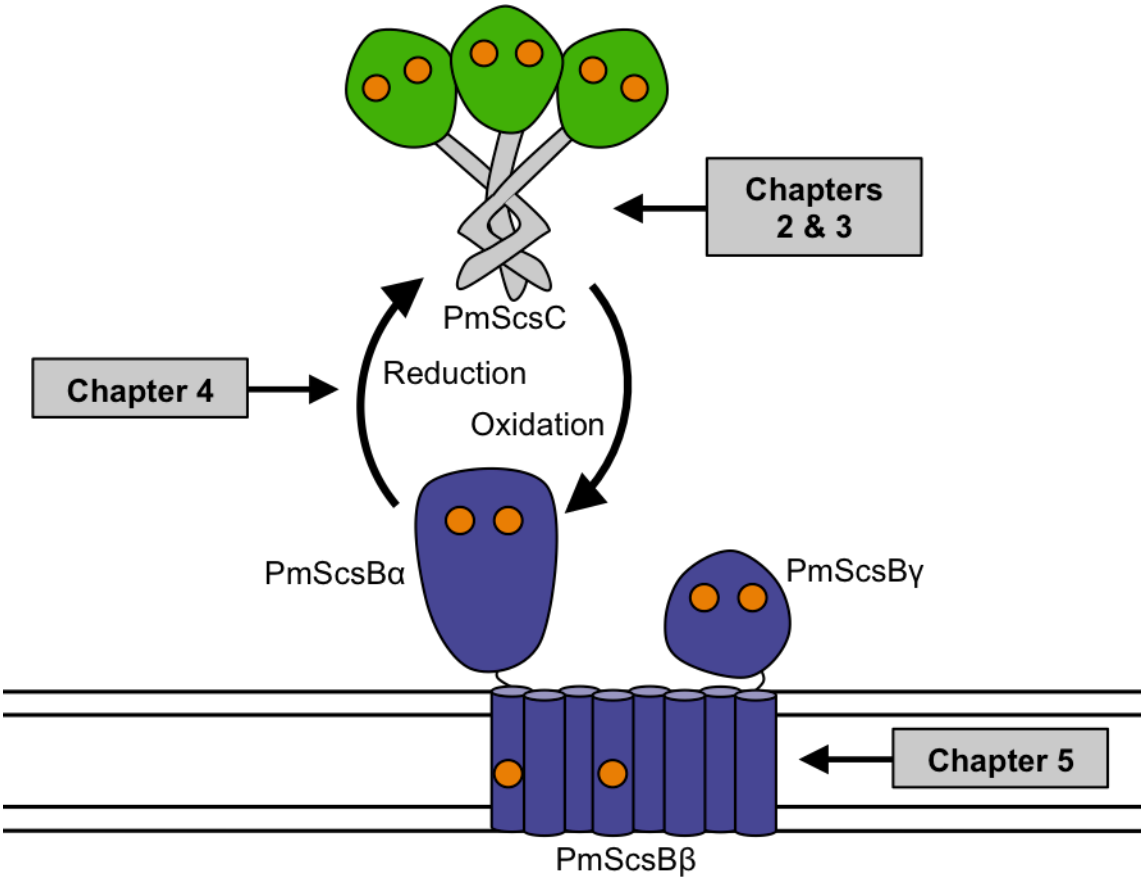


Figure 6.1: Summary of the work presented in this thesis. *Chapters 2 and 3 describe the research into PmScsC showing that it is a trimeric disulfide isomerase. Chapter 4 describes my characterisation of the interaction between PmScsC and PmScsBα, and reports the PmScsBα crystal structure that I solved. The initial outcomes from my research to develop a high-yielding expression and purification protocol for full-length PmScsB are given in Chapter 5.*

6.1 Understanding PmScsC

The first aim of this PhD thesis was to characterise PmScsC. ScsC from *S.* Typhimurium and *C. crescentus* had previously been characterised, revealing that StScsC is a monomeric oxidoreductase and CcScsC is a dimeric disulfide isomerase (118,120). As shown in Chapter 2, PmScsC is different to both StScsC and CcScsC. PmScsC functions as a disulfide isomerase but unlike CcScsC and other disulfide isomerases (39,118), PmScsC is a homotrimer. Three crystal structures of PmScsC were solved, revealing that the protein trimerises via its N-terminal helices and, like other DsbA-like proteins, has catalytic domains consisting of a thioredoxin fold and an inserted helical fold. The three crystal structures of the trimer had very different conformations. Taken together with SAXS data of the protein in solution, it became apparent that the trimeric disulfide isomerase is highly dynamic.

Variants of PmScsC were created to understand the role of specific structural elements for disulfide isomerase activity. PmScsCΔN lacks the N-terminal 41 amino acids (which form the trimerisation stalk α-helices 1 and 2) of the mature PmScsC protein. In chapter 2, this variant was functionally characterised showing that it does not form a trimer and does not function as an efficient disulfide isomerase. However, PmScsCΔN has a much higher level of dithiol oxidase activity (ability to introduce disulphide bonds) than the native protein. This confirmed that the N-terminal helices are required for trimerisation and that PmScsC needs to be trimeric to have disulfide isomerase activity. Chapter 3 reports the crystal structure of PmScsCΔN, showing that it is almost identical to the catalytic domains of the native PmScsC crystal structures. This outcome suggested that the acquired dithiol oxidase activity of PmScsCΔN is a consequence of the loss of the trimerisation stem and the associated change in oligomerisation state.

The region of the protein that facilitates the different protomer conformations in the three crystal structures is an 11-amino acid flexible linker, situated within the trimerisation stem. This linker is able to adopt different secondary structure conformations to position the catalytic domain in different orientations. To assess the role of this linker, two variants of PmScsC were created. The first variant, which is described in chapter 2, is the PmScsC RHP variant, where I replaced the flexible linker with a rigid helical peptide. The rigid helical peptide linker has been reported previously (135), and is stabilised in a helical conformation by the formation of salt bridges between glutamate and lysine resides, separated by three alanine residues. The PmScsC RHP variant remained trimeric and the

catalytic motifs were accessible. SAXS data showed that the protein had significantly reduced motion compared with the native protein. My analysis of the RHP variant showed that its disulfide isomerase activity in the scrambled RNase A assay was significantly reduced. Together these findings suggested that the flexible linker, and thus the range of motion it provides, is important for the disulfide isomerase activity of PmScsC.

Chapter 3 describes the second PmScsC linker variant, PmScsCΔLinker. In this variant, the residues of the 11-amino acid flexible linker are deleted (and not replaced) from the protein sequence. The crystal structure of PmScsCΔLinker is trimeric; however, SAXS analysis revealed that in solution the protein exists in an equilibrium between different oligomeric states. At low concentrations the protein is primarily monomeric. As the concentration increases a low affinity dimer is favoured over the trimeric form. This favouring of the dimer form over the trimer form is probably due to unfavourable interactions between catalytic domains. In the trimeric PmScsCΔLinker crystal structure the catalytic domains are packed very closely together, as a result of the shortened trimerisation stem. These findings suggested that the flexible linker is not only important for the dynamics of PmScsC but also the formation of the trimer, as it acts as a spacer between the catalytic domains. As oligomerisation is important for activity of disulfide isomerases (41,123,175), it was unsurprising that the PmScsCΔLinker variant was inactive in the disulfide isomerase assay. Unlike PmScsCΔN, this variant is also a poor dithiol oxidase despite being predominantly monomeric at the concentration used in the assay. One possible explanation for this result is that the trimerisation stem somehow interferes with peptide substrate binding.

Given its function as an efficient disulfide isomerase and its highly dynamic nature it is possible that PmScsC could be useful in the biotechnology industry, for example to assist in the folding of recombinant proteins in *E. coli* expression systems. An *E. coli* strain expressing cytoplasmic EcDsbC (SHuffle®) is already commercially available to assist folding of multi-disulfide bonded proteins (240). Given PmScsC's trimeric and highly dynamic nature, it may be able to refold different or a greater variety of substrates. Future work towards this area of research could involve trial co-expressions of PmScsC with multi-disulfide bonded proteins that do not fold well in typical *E. coli* expression systems and comparison with the SHuffle® system.

Further modifications of the trimerisation stem and flexible linker to alter PmScsC function could also be explored. The flexible linker could be replaced entirely with alanine residues or alanine scanning could be used to determine which specific amino acids are important in the linker. The flexible linker and/or trimerisation stem could also be attached to other Dsb-like catalytic domains to determine whether this converts them into trimeric proteins and disulfide isomerases.

Further work could focus on understanding the specific *in vivo* role of PmScsC. In chapter 2 it was shown that deletion of *scsC* from *P. mirabilis* results in reduced swarming motility under copper stress and successful complementation occurred only when both catalytic cysteines were available and the protein was trimeric. This suggests that PmScsC needs to be a trimer and has redox activity *in vivo* but we do not know the precise role the protein plays in the periplasm or how it contributes to copper resistance. We do not know whether PmScsC directly binds copper or shuffles incorrect copper-catalysed disulfide bonds in substrate proteins or has some other role. Previous work in *S.* Typhimurium has suggested that copper is required to turn the *scs* operon on (112) but nothing else is known about the interaction between copper and the Scs proteins. Identifying PmScsC substrates and assessing its copper binding ability will assist in understanding its specific function and give insight into its role in copper resistance. Substrates could be predicted using bioinformatic tools such as STRING (241) and then these substrates could be confirmed *in vitro* (by recombinantly expressing and purifying the proteins and performing interaction assays) or *in vivo* (by analysing the abundance and function of these proteins in *scsC* deletion strains of *P. mirabilis*). Copper binding can be determined by the use of size exclusion chromatography with spectrophotometric assays (242,243) and the *in vitro* binding affinity can be quantified by analysing the competition between PmScsC and copper chelators. The results from these studies may also help explain why PmScsC is trimeric whereas ScsC proteins from other bacterial species are monomeric and why PmScsC is highly dynamic whereas other disulfide isomerases do not exhibit the same dynamic nature. It would also be interesting to investigate if the expression of ScsC and other Scs proteins in *P. mirabilis* is dependent on the presence of copper, as was found for *S.* Typhimurium (112).

6.2 PmScsBα and its role in PmScsC activation

Work by others on ScsB from *C. crescentus* showed that it maintains ScsC in the reduced state and predicted that the N-terminal periplasmic domain of ScsB (ScsBα) was responsible for directly interacting with ScsC in that organism (118). These results suggested that PmScsC and PmScsB may form a redox relay. My research in chapter 4 of this thesis confirmed that disulfide exchange occurs *in vitro* between reduced ScsBα and oxidised ScsC from *P. mirabilis*, converting PmScsC to the reduced form and restoring its isomerase activity.

The crystal structure of PmScsBα, the first ScsBα structure to be solved, was also reported in chapter 4. This structure revealed that the PmScsBα domain consists of two immunoglobulin-like folds connected by a single α-helix. This arrangement differs from the α domain of EcDsbD, which has just one immunoglobulin-like fold. The catalytic cysteines are present in the more N-terminal of PmScsBα's two immunoglobulin-like folds. In both PmScsB and EcDsbD both cysteines are required to reduce their substrates PmScsC and EcDsbC, respectively.

Chapter 4 also reported the low-resolution structure of the complex between PmScsBα and PmScsC. The interaction between the two is transient so to generate a stable complex for structural studies, one catalytic cysteine in each of the two proteins was mutated. This allowed formation of a stable intermolecular bond between the proteins. SANS analysis was used to determine the low-resolution model for the complex, revealing that both immunoglobulin-like folds of PmScsBα bind to a single PmScsC protomer within the trimer. This interaction differs from that between dimeric disulfide isomerase EcDsbC and its redox partner EcDsbDα, in which EcDsbDα interacts with both protomers of the EcDsbC homodimer (47,61).

The PmScsC-PmScsBα SANS model suggests that PmScsBα does not interact with two of the three protomers or with the trimerisation stem of PmScsC. This finding suggests that the mode of interaction we observed could be conserved between ScsBα and non-trimeric ScsC proteins, like those from *S.* Typhimurium (monomeric) and *C. crescentus* (proposed dimeric). Future work is needed to confirm this hypothesis by investigating the interaction between PmScsBα and monomeric PmScsCΔN or between ScsB and ScsC proteins from other bacterial species. The disulfide exchange between the two proteins could be assessed using a gel-shift assay with AMS, which binds to free thiols. If the interaction

exists, an intermolecular disulfide bond stabilised complex between the proteins could be created, as described previously, and used for structural studies.

It should be noted that PmScsBα is only one domain of a 3-domain protein, the other two domains being the transmembrane domain and the C-terminal periplasmic domain. Therefore it is possible that the interaction mode between PmScsC and PmScsBα may differ in the presence of the additional domains. Future work should follow up on the interaction between PmScsC and the full-length transmembrane protein PmScsB. As described previously, cysteine variants of the proteins could be produced and used to form a stable complex, which would allow the stoichiometry of the interaction to be assessed and could be used in structural studies.

6.3 Production of PmScsB using a superfolder GFP fusion

This PhD work also aimed to develop an expression and purification protocol for the production of PmScsB. Such a protocol could be applied in future work on the characterisation of the full-length protein and its interaction with PmScsC. Chapter 5 described the work towards this aim and showed that PmScsB can be recombinantly expressed in *E. coli* with a superfolder GFP fusion tag and purified using IMAC, TEV protease cleavage and SEC. Future work can build upon this initial expression and purification protocol I developed to increase the yield and stability of full-length PmScsB. The expression of PmScsB could be further optimised by screening different concentrations of L-rhamnose in the expression media. Buffer and detergent screening using the unfolding assay with CPM dye (228) may identify conditions that increase the stability of PmScsB for use throughout purification.

Given that ScsB is an electron transporter with two mobile periplasmic domains it is likely that PmScsB will be conformationally heterogeneous in solution and thus structural studies may be challenging. One way to reduce the conformational variability would be to form a stable inter-domain disulfide between PmScsBα and PmScsBγ, by mutating one catalytic cysteine in each domain. This PmScsB variant could then be used in crystallisation screening. Alternatively, forming a stable complex between PmScsB and PmScsC may reduce the conformational flexibility of both proteins. This stable complex could be generated by oxidising purified PmScsC C87S and PmScsB C114A together, which would form an unresolvable intermolecular disulfide bond between the two proteins. As the molecular weight of the PmScsB-PmScsC complex would be ~150 kDa, it would be worth trying electron microscopy for structural studies. For this, negative stain electron microscopy could be used first to determine the homogeneity of the sample and provide low-resolution structural information. If the sample was shown to be of good quality, cryo-electron microscopy could be used to gain higher resolution information.

6.4 Other future directions

The work in this thesis has focused on two of the Scs proteins, ScsC and ScsB, but ScsA and ScsD have not been biochemically or structurally characterised in any organism, to date. Future work could involve the expression, purification and biochemical and structural characterisation of both these transmembrane proteins. It is predicted that ScsD interacts with ScsB in *S.* Typhimurium and *C. crescentus*, so this interaction could be assessed using the same AMS gel-shift approach outlined in section 4.6.8. Standard redox biochemical assays could also be used to investigate whether ScsD functions as an oxidoreductase and to determine the pK_a and redox potential values of the protein. If the full-length membrane protein ScsD proves difficult to work with, the N-terminal anchor could be removed to work on the more tractable C-terminal periplasmic domain. This should be more easily expressed and purified and may be amenable to crystallisation.

ScsA is likely to be more challenging to characterise, as it has no homology to any characterised protein. There are no predicted interaction partners but - given that it is involved in the bacterial response to oxidative stress - a good starting point would be to determine if it has antioxidant activity *in vitro*. This could be done using the metal-catalysed oxidation assay that would determine the ability of ScsA to protect DNA from oxidative damage caused by reactive oxygen species generated from iron and dithiothreitol (244). Analysis of the molecular weight of the mature protein could be performed using mass spectrometry or SDS-PAGE to determine if ScsA does have a cleavable signal sequence or if it instead forms two transmembrane helices. Structural characterisation of ScsA by X-ray crystallography is likely to be challenging given that there is no predicted soluble domain, so there would be little opportunity for crystal contacts to form. Cryo-electron microscopy is unsuitable because the molecular weight of ScsA is only ~15 kDa but this makes it a good candidate for structural characterisation by NMR.

Another topic for future study is the functional similarity and interplay between the Scs and Dsb proteins. Key research questions to be answered are:

i. Can PmScsC, StScsC or CcScsC functionally replace EcDsbC *in vivo*?
ii. Can EcDsbC replace ScsC functionally in *P. mirabilis*, *C. crescentus* or *S.* Typhimurium?
iii. In organisms, such as *P. mirabilis,* that encode DsbCs and ScsB but not DsbD, are the DsbC proteins reduced by ScsB?

iv. In organisms, such as *S.* Typhimurium, that have all Dsb and Scs proteins, can DsbC interact with ScsB and ScsC with DsbD?

These questions could be answered using the AMS gel-shift assay with purified proteins and *in vivo* experiments with *scs* and *dsb* deletion strains of bacteria. The results from these experiments would increase our understanding about the functional relatedness of and cross talk between these two similar, yet different, bacterial redox machineries.

6.5 Concluding remarks

This PhD thesis describes the characterisation of PmScsC, the crystal structure of PmScsBα, the characterisation of the interaction between these two proteins and an initial expression and purification protocol for full-length PmScsB is outlined. More broadly, this work has increased our understanding of the bacterial suppressor of copper sensitivity proteins. Nevertheless, there is still much to discover and key questions remain to be answered (Figure 6.2). The specific role that all four Scs proteins play in the bacterial response to copper and oxidative stress is unknown and the reason for such diversity in ScsC proteins is not understood. The biochemical function and structure of ScsA and ScsD is also unknown and it is unclear as to what cross talk or redundancy there is, if any, between the Dsb and Scs pathways. Future work could address these knowledge gaps to build upon the findings outlined in this thesis.

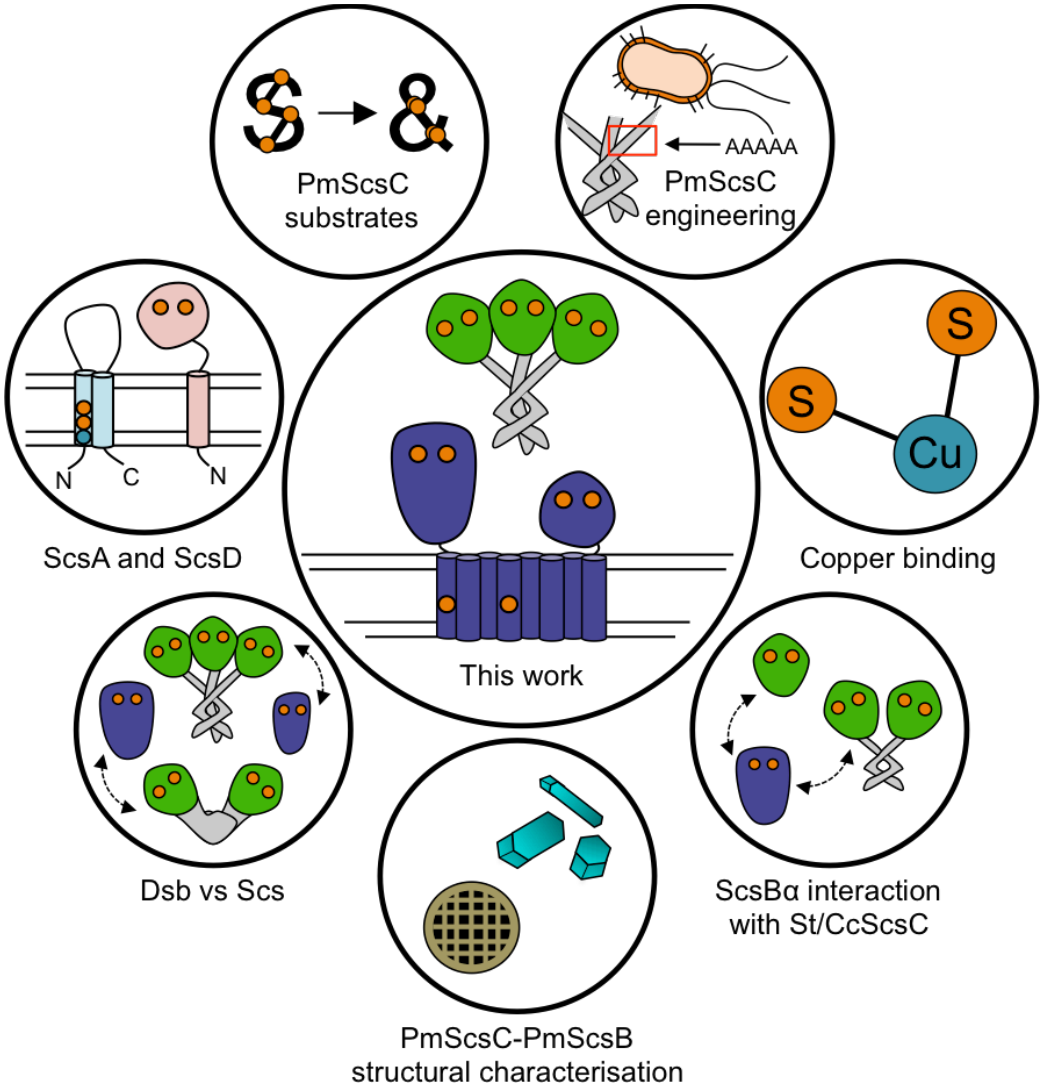


Figure 6.2: Summary of the future directions that stem from the work outlined in this thesis.

7. List of References

1. Bardwell, J. C., McGovern, K., and Beckwith, J. (1991) Identification of a protein required for disulfide bond formation *in vivo*. *Cell* **67**, 581-589
2. Peek, J. A., and Taylor, R. K. (1992) Characterization of a periplasmic thiol:disulfide interchange protein required for the functional maturation of secreted virulence factors of *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 6210-6214
3. Bardwell, J. C., Lee, J. O., Jander, G., Martin, M., Belin, D., and Beckwith, J. (1993) A pathway for disulfide bond formation *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1038-1042
4. Missiakas, D., Georgopoulos, C., and Raina, S. (1994) The *Escherichia coli dsbC* (*xprA*) gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J.* **13**, 2013-2020
5. Shevchik, V. E., Condemine, G., and Robert-Baudouy, J. (1994) Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity. *EMBO J.* **13**, 2007-2012
6. Missiakas, D., Schwager, F., and Raina, S. (1995) Identification and characterisation of a new disulfide isomerase-like protein (DsbD) in *Eschericia coli*. *EMBO J.* **14**, 3415-3424
7. Martin, J. L., Bardwell, J. C., and Kuriyan, J. (1993) Crystal structure of the DsbA protein required for disulfide bond formation *in vivo*. *Nature* **365**, 464-468
8. Nelson, J. W., and Creighton, T. E. (1994) Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation *in vivo*. *Biochemistry* **33**, 5974-5983
9. Zapun, A., Bardwell, J., and Creighton, T. E. (1993) The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation *in vivo*. *Biochemistry* **32**, 5083-5092
10. Huber-Wunderlich, M., and Glockshuber, R. (1998) A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold. Des.* **3**, 161-171
11. Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G. W., Holst, M. J., McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112-128
12. Fernandes, P. A., and Ramos, M. J. (2004) Theoretical insights into the mechanism for thiol/disulfide exchange. *Chem. Eur. J.* **10**, 257-266

13. Bach, R. D., Dmitrenko, O., and Thorpe, C. (2008) Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J. Org. Chem.* **73**, 12-21
14. Frech, C., Wunderlich, M., Glockshuber, R., and Schmid, F. X. (1996) Preferential binding of an unfolded protein to DsbA. *EMBO J.* **15**, 392-398
15. Kadokura, H., and Beckwith, J. (2009) Detecting folding intermediates of a protein as it passes through the bacterial translocation channel. *Cell* **138**, 1164-1173
16. Heras, B., Shouldice, S. R., Totsika, M., Scanlon, M. J., Schembri, M. A., and Martin, J. L. (2009) DSB proteins and bacterial pathogenicity. *Nat. Rev. Microbiol.* **7**, 215-225
17. Hiniker, A., and Bardwell, J. C. (2004) *In vivo* substrate specificity of periplasmic disulfide oxidoreductases. *J. Biol. Chem.* **279**, 12967-12973
18. Miki, T., Okada, N., Kim, Y., Abe, A., and Danbara, H. (2008) DsbA directs efficient expression of outer membrane secretin EscC of the enteropathogenic *Escherichia coli* type III secretion apparatus. *Microb. Pathog.* **44**, 151-158
19. Jander, G., Martin, N. L., and Beckwith, J. (1994) Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. *EMBO J.* **13**, 5121-5127
20. Guilhot, C., Jander, G., Martin, N. L., and Beckwith, J. (1995) Evidence that the pathway of disulfide bond formation in *Escherichia coli* involves interactions between the cysteines of DsbB and DsbA. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 9895-9899
21. Kishigami, S., and Ito, K. (1996) Roles of cysteine residues of DsbB in its activity to reoxidize DsbA, the protein disulfide catalyst of *Escherichia coli*. *Genes Cells* **1**, 201-208
22. Kobayashi, T., and Ito, K. (1999) Respiratory chain strongly oxidizes the CXXC motif of DsbB in the *Escherichia coli* disulfide bond formation pathway. *EMBO J.* **18**, 1192-1198
23. Bader, M. W., Xie, T., Yu, C. A., and Bardwell, J. C. (2000) Disulfide bonds are generated by quinone reduction. *J. Biol. Chem.* **275**, 26082-26088
24. Bader, M., Muse, W., Ballou, D. P., Gassner, C., and Bardwell, J. C. A. (1999) Oxidative protein folding is driven by the electron transport system. *Cell* **96**, 217-227
25. Inaba, K., Murakami, S., Suzuki, M., Nakagawa, A., Yamashita, E., Okada, K., and Ito, K. (2006) Crystal structure of the DsbB-DsbA complex reveals a mechanism of disulfide bond generation. *Cell* **127**, 789-801

26. Inaba, K., Murakami, S., Nakagawa, A., Iida, H., Kinjo, M., Ito, K., and Suzuki, M. (2009) Dynamic nature of disulphide bond formation catalysts revealed by crystal structures of DsbB. *EMBO J.* **28**, 770-791

27. Zhou, Y., Cierpicki, T., Jimenez, R. H., Lukasik, S. M., Ellena, J. F., Cafiso, D. S., Kadokura, H., Beckwith, J., and Bushweller, J. H. (2008) NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Mol. Cell* **31**, 896-908

28. Inaba, K., Takahashi, Y. H., Fujieda, N., Kano, K., Miyoshi, H., and Ito, K. (2004) DsbB elicits a red-shift of bound ubiquinone during the catalysis of DsbA oxidation. *J. Biol. Chem.* **279**, 6761-6768

29. Inaba, K., Takahashi, Y. H., Ito, K., and Hayashi, S. (2006) Critical role of a thiolate-quinone charge transfer complex and its adduct form in *de novo* disulfide bond generation by DsbB. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 287-292

30. Totsika, M., Heras, B., Wurpel, D. J., and Schembri, M. A. (2009) Characterization of two homologous disulfide bond systems involved in virulence factor biogenesis in uropathogenic *Escherichia coli* CFT073. *J. Bacteriol.* **191**, 3901-3908

31. Adams, L. A., Sharma, P., Mohanty, B., Ilyichova, O. V., Mulcair, M. D., Williams, M. L., Gleeson, E. C., Totsika, M., Doak, B. C., Caria, S., Rimmer, K., Horne, J., Shouldice, S. R., Vazirani, M., Headey, S. J., Plumb, B. R., Martin, J. L., Heras, B., Simpson, J. S., and Scanlon, M. J. (2015) Application of fragment-based screening to the design of inhibitors of *Escherichia coli* DsbA. *Angew. Chem. Int. Ed. Engl.* **54**, 2179-2184

32. Duprez, W., Premkumar, L., Halili, M. A., Lindahl, F., Reid, R. C., Fairlie, D. P., and Martin, J. L. (2015) Peptide inhibitors of the *Escherichia coli* DsbA oxidative machinery essential for bacterial virulence. *J. Med. Chem.* **58**, 577-587

33. Duprez, W., Bachu, P., Stoermer, M. J., Tay, S., McMahon, R. M., Fairlie, D. P., and Martin, J. L. (2015) Virtual screening of peptide and peptidomimetic fragments targeted to inhibit bacterial dithiol oxidase DsbA. *PLOS ONE* **10**, e0133805

34. Landeta, C., Blazyk, J. L., Hatahet, F., Meehan, B. M., Eser, M., Myrick, A., Bronstain, L., Minami, S., Arnold, H., Ke, N., Rubin, E. J., Furie, B. C., Furie, B., Beckwith, J., Dutton, R., and Boyd, D. (2015) Compounds targeting disulfide bond forming enzyme DsbB of Gram-negative bacteria. *Nat. Chem. Biol.* **11**, 292-298

35. Fruh, V., Zhou, Y., Chen, D., Loch, C., Ab, E., Grinkova, Y. N., Verheij, H., Sligar, S. G., Bushweller, J. H., and Siegal, G. (2010) Application of fragment-based drug

discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. *Chem. Biol.* **17**, 881-891

36. Halili, M. A., Bachu, P., Lindahl, F., Bechara, C., Mohanty, B., Reid, R. C., Scanlon, M. J., Robinson, C. V., Fairlie, D. P., and Martin, J. L. (2015) Small molecule inhibitors of disulfide bond formation by the bacterial DsbA-DsbB dual enzyme system. *ACS Chem. Biol.* **10**, 957-964
37. Totsika, M., Vagenas, D., Paxman, J. J., Wang, G., Dhouib, R., Sharma, P., Martin, J. L., Scanlon, M. J., and Heras, B. (2018) Inhibition of diverse DsbA enzymes in multi-DsbA encoding pathogens. *Antioxid. Redox Signal.* **29**, 653-666
38. Mohanty, B., Rimmer, K., McMahon, R. M., Headey, S. J., Vazirani, M., Shouldice, S. R., Coincon, M., Tay, S., Morton, C. J., Simpson, J. S., Martin, J. L., and Scanlon, M. J. (2017) Fragment library screening identifies hits that bind to the non-catalytic surface of *Pseudomonas aeruginosa* DsbA1. *PLOS ONE* **12**, e0173436
39. McCarthy, A. A., Haebel, P. W., Torronen, A., Rybin, V., Baker, E. N., and Metcalf, P. (2000) Crystal structure of the protein disulfide isomerase, DsbC, from *Escherichia coli*. *Nat. Struct. Biol.* **7**, 196-199
40. Sone, M., Akiyama, Y., and Ito, K. (1997) Differential *in vivo* roles played by DsbA and DsbC in the formation of protein disulfide bonds. *J. Biol. Chem.* **272**, 10349-10352
41. Sun, X. X., and Wang, C. C. (2000) The N-terminal sequence (residues 1-65) is essential for dimerization, activities, and peptide binding of *Escherichia coli* DsbC. *J. Biol. Chem.* **275**, 22743-22749
42. Zapun, A., Missiakas, D., Raina, S., and Creighton, T. E. (1995) Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry* **34**, 5075-5089
43. Zhao, Z., Peng, Y., Hao, S. F., Zeng, Z. H., and Wang, C. C. (2003) Dimerization by domain hybridization bestows chaperone and isomerase activities. *J. Biol. Chem.* **278**, 43292-43298
44. Goldstone, D., Haebel, P. W., Katzen, F., Bader, M. W., Bardwell, J. C., Beckwith, J., and Metcalf, P. (2001) DsbC activation by the N-terminal domain of DsbD. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9551-9556
45. Bader, M. W., Hiniker, A., Regeimbal, J., Goldstone, D., Haebel, P. W., Riemer, J., Metcalf, P., and Bardwell, J. (2001) Turning a disulfide isomerase into an oxidase: DsbC mutants that imitate DsbA. *EMBO J.* **20**, 1555-1562

46. Chen, J., Song, J., Zhang, S., Wang, Y., Cui, D., and Wang, C. (1999) Chaperone activity of DsbC. *J. Biol. Chem.* **274**, 19601-19605

47. Haebel, P. W., Goldstone, D., Katzen, F., Beckwith, J., and Metcalf, P. (2002) The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbDα complex. *EMBO J.* **21**, 4774-4784

48. Rietsch, A., Bessette, P., Georgiou, G., and Beckwith, J. (1997) Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J. Bacteriol.* **179**, 6602-6608

49. Shouldice, S. R., Cho, S. H., Boyd, D., Heras, B., Eser, M., Beckwith, J., Riggs, P., Martin, J. L., and Berkmen, M. (2010) *In vivo* oxidative protein folding can be facilitated by oxidation-reduction cycling. *Mol. Microbiol.* **75**, 13-28

50. Maskos, K., Huber-Wunderlich, M., and Glockshuber, R. (2003) DsbA and DsbC-catalyzed oxidative folding of proteins with complex disulfide bridge patterns *in vitro* and *in vivo*. *J. Mol. Biol.* **325**, 495-513

51. Pan, J. L., and Bardwell, J. C. (2006) The origami of thioredoxin-like folds. *Protein Sci.* **15**, 2217-2227

52. Vertommen, D., Depuydt, M., Pan, J., Leverrier, P., Knoops, L., Szikora, J. P., Messens, J., Bardwell, J. C., and Collet, J. F. (2008) The disulphide isomerase DsbC cooperates with the oxidase DsbA in a DsbD-independent manner. *Mol. Microbiol.* **67**, 336-349

53. Rietsch, A., Belin, D., Martin, N., and Beckwith, J. (1996) An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13048-13053

54. Hiniker, A., Collet, J. F., and Bardwell, J. C. (2005) Copper stress causes an *in vivo* requirement for the *Escherichia coli* disulfide isomerase DsbC. *J. Biol. Chem.* **280**, 33785-33791

55. Denoncin, K., Vertommen, D., Arts, I. S., Goemans, C. V., Rahuel-Clermont, S., Messens, J., and Collet, J. F. (2014) A new role for *Escherichia coli* DsbC protein in protection against oxidative stress. *J. Biol. Chem.* **289**, 12356-12364

56. Chung, J., Chen, T., and Missiakas, D. (2000) Transfer of electrons across the cytoplasmic membrane by DsbD, a membrane protein involved in thiol-disulphide exchange and protein folding in the bacterial periplasm. *Mol. Microbiol.* **35**, 1099-1109

57. Stewart, E. J., Katzen, F., and Beckwith, J. (1999) Six conserved cysteines of the membrane protein DsbD are required for the transfer of electrons from the cytoplasm to the periplasm of *Escherichia coli*. *EMBO J.* **18**, 5963-5971

58. Gordon, E. H. J., Page, M. D., Willis, A. C., and Ferguson, S. J. (2000) *Escherichia coli* DipZ: anatomy of a transmembrane protein disulphide reductase in which three pairs of cysteine residues, one in each of three domains, contribute to function. *Mol. Microbiol.* **35**, 1360-1374

59. Katzen, F., and Beckwith, J. (2000) Transmembrane electron transfer by the membrane protein DsbD occurs via a disulfide bond cascade. *Cell* **103**, 769-779

60. Collet, J. F., Riemer, J., Bader, M. W., and Bardwell, J. C. (2002) Reconstitution of a disulfide isomerization system. *J. Biol. Chem.* **277**, 26886-26892

61. Goulding, C. W., Sawaya, M. R., Parseghian, A., Lim, V., Eisenberg, D., and Missiakas, D. (2002) Thiol-disulfide exchange in an immunoglobulin-like fold: Structure of the N-terminal domain of DsbD. *Biochemistry* **41**, 6920-6927

62. Stirnimann, C. U., Rozhkova, A., Grauschopf, U., Bockmann, R. A., Glockshuber, R., Capitani, G., and Grutter, M. G. (2006) High-resolution structures of *Escherichia coli* cDsbD in different redox states: A combined crystallographic, biochemical and computational study. *J. Mol. Biol.* **358**, 829-845

63. Kim, J. H., Kim, S. J., Jeong, D. G., Son, J. H., and Ryu, S. E. (2003) Crystal structure of DsbDγ reveals the mechanism of redox potential shift and substrate specificity. *FEBS Lett.* **543**, 164-169

64. Rozhkova, A., Stirnimann, C. U., Frei, P., Grauschopf, U., Brunisholz, R., Gutter, M. G., Capitani, G., and Glockshuber, R. (2004) Structural basis and kinetics of inter- and intramolecular disulfide exchange in the redox catalyst DsbD. *EMBO J.* **23**, 1709-1719

65. Cho, S. H., and Beckwith, J. (2009) Two snapshots of electron transport across the membrane: insights into the structure and function of DsbD. *J. Biol. Chem.* **284**, 11416-11424

66. Cho, S. H., Porat, A., Ye, J., and Beckwith, J. (2007) Redox-active cysteines of a membrane electron transporter DsbD show dual compartment accessibility. *EMBO J.* **26**, 3509-3520

67. Depuydt, M., Leonard, S. E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., Messens, J., Caroll, K. S., and Collet, J. F. (2009) A periplasmic reducing system protects single cysteine residues from oxidation. *Science* **326**, 1109-1111

68. Bessette, P., Cotto, J. J., Gilbert, H. F., and Georgiou, G. (1999) *In vivo* and *in vitro* function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. *J. Biol. Chem.* **274**, 7784-7792

69. Edeling, M. A., Ahuja, U., Heras, B., Thony-Meyer, L., and Martin, J. L. (2004) The acidic nature of the CcmG redox-active center is important for cytochrome c maturation in *Escherichia coli*. *J. Bacteriol.* **186**, 4030-4033

70. Heras, B., Edeling, M. A., Schirra, H. J., Raina, S., and Martin, J. L. (2004) Crystal structures of the DsbG disulfide isomerase reveal an unstable disulfide. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8876-8881

71. Shao, F., Bader, M. W., Jakob, U., and Bardwell, J. C. A. (2000) DsbG, a protein disulfide isomerase with chaperone activity. *J. Biol. Chem.* **275**, 13349-13352

72. Kouidmi, I., Alvarez, L., Collet, J. F., Cava, F., and Paradis-Bleau, C. (2018) The chaperone activity of DsbG and Spy restore peptidoglycan biosynthesis arrest in the *elyC* mutant by preventing envelope protein aggregation. *J. Bacteriol.* **200**, e00245-00218

73. Mainardi, J. L., Hugonnet, J. E., Rusconi, F., Fourgeaud, M., Dubost, L., Moumi, A. N., Delfosse, V., Mayer, C., Gutmann, L., Rice, L. B., and Arthur, M. (2007) Unexpected inhibition of peptidoglycan LD-transpeptidase from *Enterococcus faecium* by the β-lactam imipenem. *J. Biol. Chem.* **282**, 30414-30422

74. Thony-Meyer, L., Fischer, F., Kunzler, P., Ritz, D., and Hennecke, H. (1995) *Escherichia coli* genes required for cytochrome c maturation. *J. Bacteriol.* **177**, 4321-4326

75. Fabianek, R. A., Hennecke, H., and Thony-Meyer, L. (1998) The active-site cysteines of the periplasmic thioredoxin-like protein CcmG of *Escherichia coli* are important but not essential for cytochrome c maturation *in vivo*. *J. Bacteriol.* **180**, 1947-1950

76. Fabianek, R. A., Hofer, T., and Thony-Meyer, L. (1999) Characterisation of the *Escherichia coli* CcmH protein reveals new insights into the redox pathway for cytochrome c maturation. *Arch. Microbiol.* **171**, 92-100

77. Stevens, J. M., Mavridou, D. A., Hamer, R., Kritsiligkou, P., Goddard, A. D., and Ferguson, S. J. (2011) Cytochrome c biogenesis system I. *FEBS J.* **278**, 4170-4178

78. Ouyang, N., Gao, Y. G., Hu, H. Y., and Xia, Z. X. (2006) Crystal structures of *E. coli* CcmG and its mutants reveal key roles of the N-terminal β-sheet and the fingerprint region. *Proteins* **65**, 1021-1031

79. Landeta, C., Boyd, D., and Beckwith, J. (2018) Disulfide bond formation in prokaryotes. *Nat. Microbiol.* **3**, 270-280

80. Bocian-Ostrzycka, K. M., Grzeszczuk, M. J., Dziewit, L., and Jagusztyn-Krynicka, E. K. (2015) Diversity of the *Epsilonproteobacteria* Dsb (disulfide bond) systems. *Front. Microbiol.* **6**, 570

81. Dumoulin, A., Grauschopf, U., Bischoff, M., Thony-Meyer, L., and Berger-Bachi, B. (2005) *Staphylococcus aureus* DsbA is a membrane-bound lipoprotein with thiol-disulfide oxidoreductase activity. *Arch. Microbiol.* **184**, 117-128

82. Kouwen, T. R., van der Goot, A., Dorenbos, R., Winter, T., Antelmann, H., Plaisier, M. C., Quax, W. J., van Dijl, J. M., and Dubois, J. Y. (2007) Thiol-disulphide oxidoreductase modules in the low-GC Gram-positive bacteria. *Mol. Microbiol.* **64**, 984-999

83. Heras, B., Kurz, M., Jarrott, R., Shouldice, S. R., Frei, P., Robin, G., Cemazar, M., Thony-Meyer, L., Glockshuber, R., and Martin, J. L. (2008) *Staphylococcus aureus* DsbA does not have a destabilizing disulfide. A new paradigm for bacterial oxidative folding. *J. Biol. Chem.* **283**, 4261-4271

84. Kwon, A.-R., and Choi, E.-C. (2005) Role of disulfide bond of arylsulfate sulfotransferase in catalytic activity. *Arch. Pharm. Res.* **28**, 561-565

85. Grimshaw, J. P., Stirnimann, C. U., Brozzo, M. S., Malojcic, G., Grutter, M. G., Capitani, G., and Glockshuber, R. (2008) DsbL and DsbI form a specific dithiol oxidase system for periplasmic arylsulfate sulfotransferase in uropathogenic *Escherichia coli*. *J. Mol. Biol.* **380**, 667-680

86. Lin, D., Kim, B., and Slauch, J. M. (2009) DsbL and DsbI contribute to periplasmic disulfide bond formation in *Salmonella enterica* serovar Typhimurium. *Microbiology* **155**, 4014-4024

87. Heras, B., Totsika, M., Jarrott, R., Shouldice, S. R., Guncar, G., Achard, M. E., Wells, T. J., Argente, M. P., McEwan, A. G., and Schembri, M. A. (2010) Structural and functional characterization of three DsbA paralogues from *Salmonella enterica* serovar Typhimurium. *J. Biol. Chem.* **285**, 18423-18432

88. Djoko, K. Y., Ong, C. L., Walker, M. J., and McEwan, A. G. (2015) The role of copper and zinc toxicity in innate immune defense against bacterial pathogens. *J. Biol. Chem.* **290**, 18954-18961

89. Crocker, A., Lee, C., Aboko-Cole, G., and Durham, C. (1992) Interaction of nutrition and infection: Effect of copper deficiency on resistance to *Trypanosoma lewisi*. *J. Natl. Med. Assoc.* **84**, 697-706

90. Jones, D. G., and Suttle, N. F. (1983) The effect of copper deficiency on the resistance of mice to infection with *Pasteurella haemolytica*. *J. Comp. Pathol.* **93**, 143-149

91. Newberne, P. M., Hunt, C. E., and Young, V. R. (1968) The role of diet and the reticuloendothelial system in the response of rats to *Salmonella* Typhimurium infection. *Br. J. Exp. Pathol.* **49**, 448-457

92. Subashchandrabose, S., Hazen, T. H., Brumbaugh, A. R., Himpsl, S. D., Smith, S. N., Ernst, R. D., Rasko, D. A., and Mobley, H. L. (2014) Host-specific induction of *Escherichia coli* fitness genes during human urinary infection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18327-18332

93. Wolschendorf, F., Ackart, D., Shrestha, T. B., Hascall-Dove, L., Nolan, S., Lamichhane, G., Wang, Y., Bossmann, S. H., Basaraba, R. J., and Niederweis, M. (2011) Copper resistance is essential for virulence of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1621-1626

94. Hodgkinson, V., and Petris, M. J. (2012) Copper homeostasis at the host-pathogen interface. *J. Biol. Chem.* **287**, 13549-13555

95. Achard, M. E., Stafford, S. L., Bokil, N. J., Chartres, J., Bernhardt, P. V., Schembri, M. A., Sweet, M. J., and McEwan, A. G. (2012) Copper redistribution in murine macrophages in response to *Salmonella* infection. *Biochem. J.* **444**, 51-57

96. White, C., Lee, J., Kambe, T., Fritsche, K., and Petris, M. J. (2009) A role for the ATP7A copper-transporting ATPase in macrophage bactericidal activity. *J. Biol. Chem.* **284**, 33949-33956

97. Hyre, A. N., Kavanagh, K., Kock, N. D., Donati, G. L., and Subashchandrabose, S. (2017) Copper is a host effector mobilized to urine during urinary tract infection to impair bacterial colonization. *Infect. Immun.* **85**, e01041-01016

98. Hurme, R., Namork, E., Nurmiaho-Lassila, E. L., and Rhen, M. (1994) Intermediate filament-like network formed *in vitro* by a bacterial coiled coil protein. *J. Biol. Chem.* **269**, 10675-10682

99. Braymer, J. J., and Giedroc, D. P. (2014) Recent developments in copper and zinc homeostasis in bacterial pathogens. *Curr. Opin. Chem. Biol.* **19**, 59-66

100. Macomber, L., and Imlay, J. A. (2009) The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8344-8349

101. Outten, F. W., Outten, C. E., Hale, J., and O'Halloran, T. V. (2000) Transcriptional activation of an *Escherichia coli* copper efflux regulon by the chromosomal MerR homologue, cueR. *J. Biol. Chem.* **275**, 31024-31029

102. Kim, J. S., Kim, M. H., Joe, M. H., Song, S. S., Lee, I. S., and Choi, S. Y. (2002) The *sctR* of *Salmonella enterica* serova Typhimurium encoding a homologue of MerR protein is involved in the copper-responsive regulation of *cuiD*. *FEMS Microbiol. Lett.* **210**, 99-103

103. Stoyanov, J. V., Hobman, J. L., and Brown, N. L. (2001) CueR (Ybbl) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. *Mol. Microbiol.* **39**, 502-511

104. Grass, G., and Rensing, C. (2001) CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **286**, 902-908

105. Singh, S. K., Roberts, S. A., McDevitt, S. F., Weichsel, A., Wildner, G. F., Grass, G. B., Rensing, C., and Montfort, W. R. (2011) Crystal structures of multicopper oxidase CueO bound to copper(I) and silver(I): functional role of a methionine-rich sequence. *J. Biol. Chem.* **286**, 37849-37857

106. Osman, D., Waldron, K. J., Denton, H., Taylor, C. M., Grant, A. J., Mastroeni, P., Robinson, N. J., and Cavet, J. S. (2010) Copper homeostasis in *Salmonella* is atypical and copper-CueP is a major periplasmic metal complex. *J. Biol. Chem.* **285**, 25259-25268

107. Djoko, K. Y., Franiek, J. A., Edwards, J. L., Falsetta, M. L., Kidd, S. P., Potter, A. J., Chen, N. H., Apicella, M. A., Jennings, M. P., and McEwan, A. G. (2012) Phenotypic characterization of a copA mutant of *Neisseria gonorrhoeae* identifies a link between copper and nitrosative stress. *Infect. Immun.* **80**, 1065-1071

108. Shafeeq, S., Yesilkaya, H., Kloosterman, T. G., Narayanan, G., Wandel, M., Andrew, P. W., Kuipers, O. P., and Morrissey, J. A. (2011) The cop operon is required for copper homeostasis and contributes to virulence in *Streptococcus pneumoniae*. *Mol. Microbiol.* **81**, 1255-1270

109. Outten, F. W., Huffman, D. L., Hale, J. A., and O'Halloran, T. V. (2001) The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*. *J. Biol. Chem.* **276**, 30670-30677

110. Franke, S., Grass, G., Rensing, C., and Nies, D. H. (2003) Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J. Bacteriol.* **185**, 3804-3812

111. Gupta, S. D., Wu, H. C., and Rick, P. D. (1997) A *Salmonella* Typhimurium genetic locus which confers copper tolerance on copper-sensitive mutants of *Escherichia coli*. *J. Bacteriol.* **179**, 4977-4984

112. Lopez, C., Checa, S. K., and Soncini, F. C. (2018) CpxR/CpxA-controls *scsABCD* transcription to counteract copper and oxidative stress in *Salmonella* Typhimurium. *J. Bacteriol.* **200**, e00126-00118

113. Kershaw, C. J., Brown, N. L., Constantinidou, C., Patel, M. D., and Hobman, J. L. (2005) The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. *Microbiology* **151**, 1187-1198

114. Pezza, A., Pontel, L. B., Lopez, C., and Soncini, F. C. (2016) Compartment and signal-specific codependence in the transcriptional control of *Salmonella* periplasmic copper homeostasis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11573-11578

115. Pontel, L. B., Scampoli, N. L., Porwollik, S., Checa, S. K., McClelland, M., and Soncini, F. C. (2014) Identification of a *Salmonella* ancillary copper detoxification mechanism by a comparative analysis of the genome-wide transcriptional response to copper and zinc excess. *Microbiology* **160**, 1659-1669

116. Yamamoto, K., and Ishihama, A. (2005) Transcriptional response of *Escherichia coli* to external copper. *Mol. Microbiol.* **56**, 215-227

117. Anwar, N., Sem, X. H., and Rhen, M. (2013) Oxidoreductases that act as conditional virulence suppressors in *Salmonella enterica* serovar Typhimurium. *PLOS ONE* **8**, e64948

118. Cho, S. H., Parsonage, D., Thurston, C., Dutton, R. J., Poole, L. B., Collet, J. F., and Beckwith, J. (2012) A new family of membrane electron transporters and its substrates, including a new cell envelope peroxiredoxin, reveal a broadened reductive capacity of the oxidative bacterial cell envelope. *mBio* **3**, 1-11

119. Verbrugghe, E., Dhaenens, M., Leyman, B., Boyen, F., Shearer, N., Van Parys, A., Haesendonck, R., Bert, W., Favoreel, H., Deforce, D., Thompson, A., Haesebrouck, F., and Pasmans, F. (2016) Host stress drives *Salmonella* recrudescence. *Sci. Rep.* **6**, 20849

120. Shepherd, M., Heras, B., Achard, M. E., King, G. J., Argente, M. P., Kurth, F., Taylor, S. L., Howard, M. J., King, N. P., Schembri, M. A., and McEwan, A. G. (2013) Structural and functional characterization of ScsC, a periplasmic thioredoxin-like protein from *Salmonella enterica* serovar Typhimurium. *Antioxid. Redox Signal.* **19**, 1494-1506

121. van Straaten, M., Missiakas, D., Raina, S., and Darby, N. J. (1998) The functional properties of DsbG, a thiol-disulfide oxidoreductase from the periplasm of *Escherichia coli*. *FEBS Lett.* **428**, 255-258

122. Ren, G., Stephan, D., Xu, Z., Zheng, Y., Tang, D., Harrison, R. S., Kurz, M., Jarrott, R., Shouldice, S. R., Hiniker, A., Martin, J. L., Heras, B., and Bardwell, J. C. (2009) Properties of the thioredoxin fold superfamily are modulated by a single amino acid residue. *J. Biol. Chem.* **284**, 10150-10159

123. Ke, H., Zhang, S., Li, J., Howlett, G. J., and Wang, C. (2006) Folding of *Escherichia coli* DsbC: Characterization of a monomeric folding intermediate. *Biochemistry* **45**

124. Edeling, M. A., Guddat, L. W., Fabianek, R. A., Thony-Meyer, L., and Martin, J. L. (2002) Structure of CcmG/DsbE at 1.14 A resolution: High-fidelity reducing activity in an indiscriminately oxidizing environment. *Structure* **10**, 973-979

125. Stirnimann, C. U., Rozhkova, A., Grauschopf, U., Grutter, M. G., Glockshuber, R., and Capitani, G. (2005) Structural basis and kinetics of DsbD-dependent cytochrome c maturation. *Structure* **13**, 985-993

126. Pearson, M. M., Sebaihia, M., Churcher, C., Quail, M. A., Seshasayee, A. S., Luscombe, N. M., Abdellah, Z., Arrosmith, C., Atkin, B., Chillingworth, T., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Walker, D., Whithead, S., Thomson, N. R., Rather, P. N., Parkhill, J., and Mobley, H. L. (2008) Complete genome sequence of uropathogenic *Proteus mirabilis*, a master of both adherence and motility. *J. Bacteriol.* **190**, 4027-4037

127. Armbruster, C. E., Mobley, H. L. T., and Pearson, M. M. (2018) Pathogenesis of *Proteus mirabilis* infection. *EcoSal Plus* **8**

128. Eriksson, S., Zbornik, J., Dahnsjö, H., Erlanson, P., Kahlmeter, O., Fritz, H., and Bauer, C.-A. (1986) The combination of Pivampicillin and Pivmecillinam versus Pivampicillin alone in the treatment of acute pyelonephritis. *Scand. J. Infect. Dis.* **18**, 431-438

129. Griffith, D. P., Musher, D. M., and Itin, C. (1976) Urease. The primary cause of infection-induced urinary stones. *Invest. Urol.* **13**, 346-350

130. Bahrani, F. K., Johnson, D. E., Robbins, D., and Mobley, H. L. T. (1991) *Proteus mirabilis* Flagella and MR/P Fimbrae: isolation, purification, N-terminal analysis, and serum antibody response following experimental urinary tract infection. *Infect. Immun.* **59**, 3574-3580

131. D'Andrea, M. M., Giani, T., De Angelis, L. H., Ciacci, N., Gniadkowski, M., Miriagou, V., Torricelli, F., and Rossolini, G. M. (2016) Draft genome sequence of *Proteus*

mirabilis NO-051/03, representative of a multidrug-resistant clone spreading in Europe and expressing the CMY-16 AmpC-Type β-lactamase. *Genome Announc.* **4**, e01702-e01715

132. Luzzaro, F., Brigante, G., D'Andrea, M. M., Pini, B., Giani, T., Mantengoli, E., Rossolini, G. M., and Toniolo, A. (2009) Spread of multidrug-resistant *Proteus mirabilis* isolates producing an AmpC-type β-lactamase: epidemiology and clinical management. *Int. J. Antimicrob. Agents* **33**, 328-333
133. Burall, L. S., Harro, J. M., Li, X., Lockatell, C. V., Himpsl, S. D., Hebel, J. R., Johnson, D. E., and Mobley, H. L. T. (2004) *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infection: Identification of 25 signature-tagged mutants attenuated at least 100-fold. *Infect. Immun.* **72**, 2922-2938
134. Kurth, F., Duprez, W., Premkumar, L., Schembri, M. A., Fairlie, D. P., and Martin, J. L. (2014) Crystal structure of the dithiol oxidase DsbA enzyme from *Proteus mirabilis* bound non-covalently to an active site peptide ligand. *J. Biol. Chem.* **289**, 19810-19822
135. Chen, X., Zaro, J. L., and Shen, W. C. (2013) Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357-1369
136. Tian, G., Kober, F. X., Lewandrowski, U., Sickmann, A., Lennarz, W. J., and Schindelin, H. (2008) The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule. *J. Biol. Chem.* **283**, 33630-33640
137. Wang, C., Li, W., Ren, J., Fang, J., Ke, H., Gong, W., Feng, W., and Wang, C. C. (2013) Structural insights into the redox-regulated dynamic conformations of human protein disulfide isomerase. *Antioxid. Redox Signal.* **19**, 36-45
138. Tian, G., Xiang, S., Noiva, R., Lennarz, W. J., and Schindelin, H. (2006) The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* **124**, 61-73
139. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207-234
140. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* (Walker, J. M. ed.), Humana Press. pp 571-607
141. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* **2**, 924-932

142. Kirby, N. M., Mudie, S. T., Hawley, A. M., Cookson, D. J., Mertens, H. D. T., Cowieson, N., and Samardzic-Boban, V. (2013) A low-background-intensity focusing small-angle X-ray scattering undulator beamline. *J. Appl. Crystallogr.* **46**, 1670-1680

143. Scatterbrain - Software for acquiring, processing and viewing SAXS/WAXS data at the Australian Synchrotron. Australian Synchrotron, Clayton, VIC

144. Orthaber, D., Bergmann, A., and Glatter, O. (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **33**, 218-225

145. Whitten, A. E., Cai, S., and Trewhella, J. (2008) MULCh: modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. *J. Appl. Crystallogr.* **41**, 222-226

146. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503

147. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V., and Svergun, D. I. (2012) New developments in the program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342-350

148. Franke, D., Jeffries, C. M., and Svergun, D. I. (2015) Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat. Methods* **12**, 419-422

149. Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217

150. McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the distributed control system: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9**, 401-406

151. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125-132

152. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204-1214

153. Vagin, A., and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 1622-1624

154. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J.,

Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-221

155. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674
156. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132
157. Dauter, Z. (2006) Estimation of anomalous signal in diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 867-876
158. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184-2195
159. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293-302
160. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235-242
161. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002-1011
162. Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 368-380
163. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C., and Womack, T. O. (2011) BUSTER version 2.10.1. Global Phasing Ltd., Cambridge, United Kingdom
164. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010)

MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21

165. Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J., and Lesk, A. M. (2006) MUSTANG: a multiple structural alignment algorithm. *Proteins* **64**, 559-574
166. Konagurthu, A. S., Reboul, C. F., Schmidberger, J. W., Irving, J. A., Lesk, A. M., Stuckey, P. J., Whisstock, J. C., and Buckle, A. M. (2010) MUSTANG-MR structural sieving server: applications in protein structural analysis and crystallography. *PLOS ONE* **5**, e10048
167. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Meth.* **9**, 671-675
168. Wunderlich, M., and Glockshuber, R. (1993) *In vivo* control of redox potential during protein folding catalyzed by bacterial protein disulfide-isomerase (DsbA). *J. Biol. Chem.* **268**, 24547-24550
169. Gilbert, H. F. (1995) Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods Enzymol.* **251**, 8-28
170. Hillson, D. A., Lambert, N., and Freedman, R. B. (1984) Formation and isomerisation of disulfide bonds in proteins: Protein disulfide-isomerase. *Methods Enzymol.* **107**, 281-294
171. Vivian, J. P., Scoullar, J., Rimmer, K., Bushell, S. R., Beddoe, T., Wilce, M. C., Byres, E., Boyle, T. P., Doak, B., Simpson, J. S., Graham, B., Heras, B., Kahler, C. M., Rossjohn, J., and Scanlon, M. J. (2009) Structure and function of the oxidoreductase DsbA1 from *Neisseria meningitidis*. *J. Mol. Biol.* **394**, 931-943
172. Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M., and Svergun, D. I. (2015) SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* **43**, D357-D363
173. Berkmen, M., Boyd, D., and Beckwith, J. (2005) The nonconsecutive disulfide bond of *Escherichia coli* phytase (AppA) renders it dependent on the protein-disulfide isomerase, DsbC. *J. Biol. Chem.* **280**, 11387-11394
174. Darby, N. J., Raina, S., and Creighton, T. E. (1998) Contributions of substrate binding to the catalytic activity of DsbC. *Biochemistry* **37**, 783-791
175. Furlong, E. J., Lo, A. W., Kurth, F., Premkumar, L., Totsika, M., Achard, M. E. S., Halili, M. A., Heras, B., Whitten, A. E., Choudhury, H. G., Schembri, M. A., and Martin, J. L. (2017) A shape-shifting redox foldase contributes to *Proteus mirabilis* copper resistance. *Nat. Commun.* **8**, 16065

176. Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E., and Martin, J. L. (2018) Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB. *J. Biol. Chem.* **293**, 5793-5805

177. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72-82

178. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501

179. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242

180. Berman, H. M., Henrick, K., and Nakamura, H. (2003) Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **10**, 980

181. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271-281

182. The PyMOL Molecular Graphics System, Version 1.6. Schrodinger, LLC, New York

183. Trewhella, J., Duff, A. P., Durand, D., Gabel, F., Guss, J. M., Hendrickson, W. A., Hura, G. L., Jacques, D. A., Kirby, N. M., Kwan, A. H., Pérez, J., Pollack, L., Ryan, T. M., Sali, A., Schneidman-Duhovny, D., Schwede, T., Svergun, D. I., Sugiyama, M., Tainer, J. A., Vachette, P., Westbrook, J., and Whitten, A. E. (2017) 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallographica Section D Structural Biology* **73**, 1-19

184. Fischer, H., Neto, M. D., Napolitano, H. B., Polikarpov, I., and Craievich, A. F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* **43**, 101-109

185. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277-1282

186. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879-2886

187. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860-864

188. Tuukkanen, A. T., Kleywegt, G. J., and Svergun, D. I. (2016) Resolution of *ab initio* shapes determined from small-angle scattering. *IUCrJ* **3**, 440-447

189. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768-773

190. Hu, S. H., Whitten, A. E., King, G. J., Jones, A., Rowland, A. F., James, D. E., and Martin, J. L. (2012) The weak complex between RhoGAP protein ARHGAP22 and signal regulatory protein 14-3-3 has 1:2 stoichiometry and a single peptide binding mode. *PLOS ONE* **7**, e41731

191. Krupp, R., Chan, C., and Missiakas, D. (2001) DsbD-catalyzed transport of electrons across the membrane of *Escherichia coli*. *J. Biol. Chem.* **276**, 3696-3701

192. Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545-549

193. Ireland, P. M., McMahon, R. M., Marshall, L. E., Halili, M., Furlong, E., Tay, S., Martin, J. L., and Sarkar-Tyson, M. (2014) Disarming *Burkholderia pseudomallei*: structural and functional characterization of a disulfide oxidoreductase (DsbA) required for virulence *in vivo*. *Antioxid Redox Signal* **20**, 606-617

194. Miki, T., Okada, N., and Danbara, H. (2004) Two periplasmic disulfide oxidoreductases, DsbA and SrgA, target outer membrane protein SpiA, a component of the *Salmonella* pathogenicity island 2 type III secretion system. *J. Biol. Chem.* **279**, 34631-34642

195. Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276-277

196. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y. M., Buso, N., and Lopez, R. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43**, W580-584

197. Pei, J., Kim, B. H., and Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295-2300

198. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr. Purif.* **25**, 8-15

199. Walden, P. M., Halili, M. A., Archbold, J. K., Lindahl, F., Fairlie, D. P., Inaba, K., and Martin, J. L. (2013) The α-proteobacteria *Wolbachia pipientis* protein disulfide

machinery has a regulatory mechanism absent in γ-proteobacteria. *PLOS ONE* **8**, e81440

200. Duff, A. P., Wilde, K. L., Rekas, A., Lake, V., and Holden, P. J. (2015) Robust high-yield methodologies for (2)H and (2)H/(15)N/(13)C labeling of proteins for structural investigations using neutron scattering and NMR. *Methods Enzymol.* **565**, 3-25
201. Sheldrick, G. M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 479-485
202. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843-844
203. Schrodinger, L. (2015) The PyMOL Molecular Graphics System, Version 1.6.
204. Gilbert, E. P., Schulz, J. C., and Noakes, T. J. (2006) 'Quokka'—the small-angle neutron scattering instrument at OPAL. *Physica B: Condensed Matter* **385-386**, 1180-1182
205. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237-1250
206. Rozhkova, A., and Glockshuber, R. (2008) Thermodynamic aspects of DsbD-mediated electron transport. *J. Mol. Biol.* **380**, 783-788
207. Katzen, F., Deshmukh, M., Daldal, F., and Beckwith, J. (2002) Evolutionary domain fusion expanded the substrate specificity of the transmembrane electron transporter DsbD. *EMBO J.* **21**, 3960-3969
208. Williamson, J. A., Cho, S. H., Ye, J., Collet, J. F., Beckwith, J. R., and Chou, J. J. (2015) Structure and multistate function of the transmembrane electron transporter CcdA. *Nat. Struct. Mol. Biol.* **22**, 809-814
209. Zhou, Y., and Bushweller, J. H. (2018) Solution structure and elevator mechanism of the membrane electron transporter CcdA. *Nat. Struct. Mol. Biol.* **25**, 163-169
210. Dobson, L., Remenyi, I., and Tusnady, G. E. (2015) CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res.* **43**, W408-412
211. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858
212. Omasits, U., Ahrens, C. H., Muller, S., and Wollscheid, B. (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30**, 884-886

213. Yin, H., and Flynn, A. D. (2016) Drugging membrane protein interactions. *Annu. Rev. Biomed. Eng.* **18**, 51-76

214. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993-996

215. Drew, D. E., von Heijne, G., Nordlund, P., and de Gier, J. W. (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Lett.* **507**, 220-224

216. Drew, D., Slotboom, D. J., Friso, G., Reda, T., Genevaux, P., Rapp, M., Meindl-Beinker, N. M., Lambert, W., Lerch, M., Daley, D. O., Van Wijk, K. J., Hirst, J., Kunji, E., and De Gier, J. W. (2005) A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci.* **14**, 2011-2017

217. Drew, D., Lerch, M., Kunji, E., Slotboom, D. J., and de Gier, J. W. (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods* **3**, 303-313

218. Kawate, T., and Gouaux, E. (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673-681

219. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G. J. (2000) Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J. Bacteriol.* **182**, 4068-4076

220. Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79-88

221. Dinh, T., and Bernhardt, T. G. (2011) Using superfolder green fluorescent protein for periplasmic protein localization studies. *J. Bacteriol.* **193**, 4984-4987

222. Wagner, S., Baars, L., Ytterberg, A. J., Klussmeier, A., Wagner, C. S., Nord, O., Nygren, P., Van Wijk, K. J., and de Gier, J. W. (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol. Cell. Proteomics* **6**, 1527-1550

223. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289-298

224. Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., and de Gier, J. W. (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14371-14376

225. Schlegel, S., Lofblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D. J., and de Gier, J. W. (2012) Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). *J. Mol. Biol.* **423**, 648-659

226. Studier, F. W. (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* **219**, 37-44

227. Anandan, A., and Vrielink, A. (2016) Detergents in membrane protein purification and crystallisation. in *The Next Generation in Membrane Protein Structure Determination* (Moraes, I. ed.), Springer International Publishing, Switzerland. pp 13-28

228. Alexandrov, A. I., Mileni, M., Chien, E. Y., Hanson, M. A., and Stevens, R. C. (2008) Microscale fluorescent thermal stability assay for membrane proteins. *Structure* **16**, 351-359

229. Sonoda, Y., Newstead, S., Hu, N. J., Alguel, Y., Nji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A. D., Byrne, B., Iwata, S., and Drew, D. (2011) Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure* **19**, 17-25

230. Denisov, I. G., and Sligar, S. G. (2017) Nanodiscs in membrane biochemistry and biophysics. *Chem. Rev.* **117**, 4669-4713

231. Shen, H. H., Lithgow, T., and Martin, L. (2013) Reconstitution of membrane proteins into model membranes: seeking better ways to retain protein activities. *Int. J. Mol. Sci.* **14**, 1589-1607

232. Cherezov, V. (2011) Lipidic cubic phase technologies for membrane protein structural studies. *Curr. Opin. Struct. Biol.* **21**, 559-566

233. Frank, J. (2017) Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nat. Protoc.* **12**, 209-212

234. Nogales, E. (2016) The development of cryo-EM into a mainstream structural biology technique. *Nat. Methods* **13**, 24-27

235. Lei, S. P., Lin, H. C., Wang, S. S., Callaway, J., and Wilcox, G. (1987) Characterization of the *Erwinia carotovora pelB* gene and its product pectate lyase. *J. Bacteriol.* **169**, 4379-4383

236. Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* **17**

237. Stetsenko, A., and Guskov, A. (2017) An overview of the top ten detergents used for membrane protein crystallization. *Crystals* **7**, 197

238. Stroud, R. M. (2011) New tools in membrane protein determination. *F1000 Biol. Rep.* **3**, 8

239. Tribet, C., Audebert, R., and Popot, J. (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15047-15050

240. Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb. Cell. Fact.* **11**, 56

241. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-452

242. Xiao, Z., Brose, J., Schimo, S., Ackland, S. M., La Fontaine, S., and Wedd, A. G. (2011) Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards. *J. Biol. Chem.* **286**, 11047-11055

243. Xiao, Z., Gottschlich, L., van der Meulen, R., Udagedara, S. R., and Wedd, A. G. (2013) Evaluation of quantitative probes for weaker Cu(i) binding sites completes a set of four capable of detecting Cu(i) affinities from nanomolar to attomolar. *Metallomics* **5**, 501-513

244. Li, J., Zhang, W.-B., Loukas, A., Lin, R.-Y., Ito, A., Zhang, L.-H., Jones, M., and McManus, D. P. (2004) Functional expression and characterization of *Echinococcus granulosus* thioredoxin peroxidase suggests a role in protection against oxidative damage. *Gene* **326**, 157-165